

The long and winding road

German scientists must persevere in the stem-cell debate, despite the occasional setback.

The German media last week trumpeted the claim that a patient in Dusseldorf with terminal heart failure has been successfully treated with adult stem cells from his own bone marrow. Although clinical researchers know that little can be learnt from a single case study, the result has already been unreasonably exploited by opponents of human embryonic stem-cell research.

The finding was published in the September issue of a German-language medical journal (M. Brehm and B. E. Strauer *Deut. Med. Wochenschr.* **132**, 1944–1948; 2007). Opponents of human embryonic stem-cell research, such as Julia Klöckner, a Christian Democrat who chairs a stem-cell committee in the German parliament, leapt on the result, claiming that the clinical success of using adult stem cells renders research on embryonic stem cells less necessary than before. They make this point just as the parliament prepares to consider whether it should modify the country's strict stem-cell laws.

The majority of scientists agree that work on both adult and embryonic sources of stem cells should run in parallel until much more is understood about their biology. But Germany is out of step with most European countries in permitting research only on human embryonic stem-cell lines that were created before January 2002, when regulations were first laid down. This situation has caused ambiguity in collaborative European Union (EU) research programmes: some partners can use new lines, but German participants could be put in jail if they did so themselves.

The past year has seen a significant shift in attitudes, however. Last November, Germany's main research-funding agency, the DFG, set the ball rolling, saying it believed that it was now time to eliminate the cut-off date. Respecting the moral dimension to which Germans are particularly sensitive, it did not suggest that German scientists should be allowed to derive their own embryonic stem-cell lines, as this involves destroying human embryos.

Response to the DFG's report has been broadly positive. A few months ago a majority in the National Ethics Council, which advises

chancellor Angela Merkel on bioethics, supported bringing the cut-off date forward to 2007. This would allow Germans to use all the cell lines involved in current EU projects. A parliamentary hearing in May also indicated that politicians of various hues would support such a relaxation in the rules. There is even wider support for any breach of the regulations to be regarded as a civil, rather than a criminal, matter.

Scientists helped to prepare the ground for this shift in opinion by patiently and thoroughly discussing with politicians and others the complex scientific issues involved. But opponents of human embryonic stem-cell research have also stepped up their campaigns. They see the recent success with adult cells as a vindication. Yet on the basis of one patient's history, it isn't even known if the recovery can be attributed to stem-cell therapy. Political leaders should be wary of taking such results at face value — especially when the stakes for human health are so high.

The reception afforded the Dusseldorf patient has disillusioned some German scientists, who feel that their painstaking efforts to get their case across have been undermined and devalued. But they should continue to promote their position whenever they can — and should adapt their strategy to match the situation in which they find themselves.

Public opinion is a fickle thing. It will not always be easy for the ad hoc group of biologists who have been working on the stem-cell issue in Germany to react to fast-moving events and make their voices clearly heard. There is no established national scientific academy in Germany to take the lead on the issue and the societies representing biologists have not quite been able to find their voice. Despite all this, researchers should persevere in their efforts to participate in — and, indeed, to lead — the stem-cell debate at every level. Eventually, their arguments will prevail. ■

"It will not always be easy for the biologists working on the stem-cell issue in Germany to make their voices clearly heard."

Genome abuse

Citizens are right to resist government pressure to expand population DNA databases.

Terrorism, crime and illegal immigration are fuelling state surveillance, and are being used to justify it to the public, who too often seem asleep to the risks of abuse. This is particularly true of national DNA databases, where in several countries there is an insidious creep to log not only serious offenders but also other classes of the population, such as immigrants and minor offenders.

So it was refreshing to see resistance articulated this month in France and the United Kingdom. Prominent French scientists led

public protests against a government bill to use DNA tests on immigrants to see whether they are related to family members already resident in the country. Such protests might seem an overreaction. Many countries already practise DNA testing of immigrants, with varying rules for use. In 1985, the first use of DNA fingerprinting for legal purposes led to a Ghanaian boy being allowed to join his family in the United Kingdom after he proved kinship (A. J. Jeffreys *et al.* *Nature* **317**, 818–819; 1985).

But the objectors are correct to argue that the French proposal, far from promoting greater fairness, is aimed at erecting another obstacle to immigration. The scientific opposition is also linked to a strong bioethical and legal tradition in France of the concept of the family as a social unit, not reduced to mere biological ties, reflecting the reality that (as in all countries) many children are not the biological offspring of

their legal father. Given this culture, there is no reason why only immigrants with a biological link should qualify for integration with their families in France. Furthermore, DNA testing of immigrants elsewhere has destroyed families by uncovering true biological relationships.

The scientists' case has enjoyed public and political support, and has embarrassed the government, which sought to defuse the controversy last week by postponing a final decision to 2009. The outcry has also thrown an overdue spotlight on issues surrounding such population databases — issues being tackled in Britain, which has the world's largest DNA fingerprint database. The National DNA Database contains samples of 4 million people or 6% of the population, and one in ten males. The Nuffield Council on Bioethics, in a landmark report this month, does a service by drawing attention to the dangers of proposals to expand the database (see <http://tinyurl.com/2upt8x>).

There is a widespread misperception, encouraged by governments and media success stories, that DNA evidence is infallible in clinching convictions or acquittals. The technology is sound, but errors or deliberate falsifications in sample taking and handling are not uncommon, and a match with a sample at the scene of a crime may amount to proof only that the person was present at some point.

Since 2003, DNA samples and fingerprints have been compulsorily taken from Britons arrested for criminal offences. But the government now proposes extending the database to include fingerprints

and DNA from anyone arrested, even for minor offences such as dropping litter. And voices within the UK government and the judiciary have suggested that the entire population should be sampled. The US government, meanwhile, is proposing to extend its database to include DNA from anyone arrested by federal agents.

The Nuffield report is right to denounce the infringements on liberty and privacy represented by such extensions as being disproportionate to any possible benefits. Suspicion of involvement in a minor offence cannot justify taking a biological sample without consent. In the United States, the largest group likely to be affected is illegal immigrants — and there is no reason to suspect this group of being more likely to engage in serious crime.

DNA fingerprints themselves contain relatively little personal information, but the biological samples are open to misuse. Although supposedly limited to direct matching of individuals for crime cases, DNA data are already used for the much less scientifically robust practices of searching for family relatives of a crime's perpetrator, and to try to reduce possible suspects to ethnic groups.

History teaches us that it is a fallacy that only those without a clear conscience need fear a knock on the door at midnight. Governments' enthusiasm for DNA databases needs to be matched by commensurate statutory protection, transparency and oversight — and vigilance by citizens. ■

Toxic alert

A method of knocking out genes in mice needs more discrimination than many have recognized.

One of the most common ways to investigate the role of a gene in human physiology is to delete its equivalent from a mouse genome and to observe the effect. The use of one enzyme in particular, the recombinase 'Cre', has revolutionized the study of gene function in mice. The technique allows researchers to introduce mutations and gene deletions in a tissue or cell type at any stage.

Hundreds of studies using this technology have been published since it was introduced more than ten years ago, shedding light on areas such as important developmental processes and the role of numerous genes in, for example, the immune or nervous systems, or in various diseases.

Briefly, it works by introducing the target DNA sequence used by the Cre enzyme, known as a *loxP* site, to either end of the gene sequence in question. By subsequently introducing the Cre enzyme, the sequence is excised. Gene targeting can be regulated by controlling where Cre is expressed or activated.

But the technology is not without its pitfalls. A number of issues have been described in a recent overview (M. Schmidt-Supprian and K. Rajewsky *Nature Immunol.* **8**, 665–668; 2007). Readers, authors and editors alike need to be alert to one particular problem: the potential toxicity of Cre expression to cells.

The induction of cell death as a consequence of Cre activity, unrelated to the targeting of any specific gene, is thought to occur when

Cre targets sites similar to *loxP* that are present in genomic DNA, thereby inducing mis-recombination and DNA damage. Most mice strains in which Cre is expressed seem to develop normally and do not show any overt signs of Cre toxicity, and it is somewhat unclear exactly under what conditions it arises. It has been suggested to result from long-term expression of high levels of the enzyme.

Regardless of the exact mechanism and circumstances, Cre toxicity is clearly a potential problem, yet in the view of some researchers it has been neglected or played down in the community. In fact, one study has systematically analysed studies using a particular Cre mouse strain and found that in more than half of the cases the appropriate control for potential Cre toxicity — the use of the same mice without the *loxP*-flanked target gene — was not included (J.-Y. Lee *et al. J. Biol. Chem.* **281**, 2649–2653; 2006). *Nature* is aware that it has in the past published papers in which such controls were lacking, although many will no doubt have been independently validated with other techniques at the time or subsequently.

It can be argued that potential toxicity due to Cre expression becomes pertinent only when the observed phenotype resulting from gene targeting involves cell death, but the complexity of biological processes probably warrants attention to the issue in all experiments. Researchers planning experiments should take into account the need for additional mice as controls. Editors at *Nature* will consider the issue and the appropriate controls with referees during the assessment of submitted papers.

No technology is without caveats, and — as the *Nature Immunology* article concludes — there will always be a degree of uncertainty with which researchers have to live. But in the interest of best scientific practice, everyone involved would be wise not to neglect the dangers and subtleties at play even in routine experiments. ■

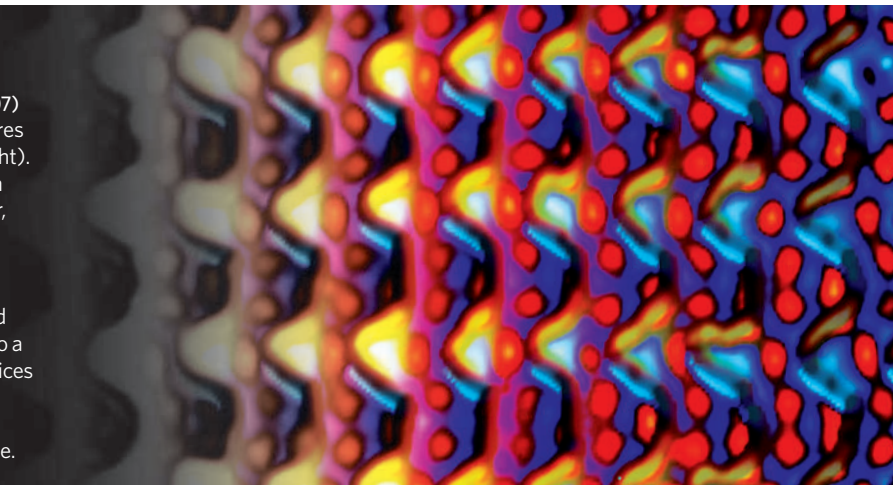
RESEARCH HIGHLIGHTS

Electric lattices

Nature Nanotechnol. doi:10.1038/nnano.2007.301 (2007)

This picture shows regularly spaced nanostructures (to the left) creating pools of electrons (to the right). The ability to form patterns of electron density on a nanometre scale might be useful for catalysis or, perhaps one day, in building spintronics devices.

Stéphane Pons and his colleagues at the CNRS basic-science agency and the University of Nancy in France began by evaporating atoms of silver and copper in a vacuum. As the atoms condensed onto a patterned gold surface, they settled into superlattices covering the whole sample. Scanning tunnelling spectroscopy revealed that the lattices localized electrons at periodically spaced sites at the surface.



LAB. PHYS. DES MATÉR./CNRS/NANCY UNIVERSITÉ

CANCER BIOLOGY

Shock strategy

Cell **130**, 1005–1018 (2007)

A protein called heat shock factor 1 (HSF1) that helps cells to handle stress seems to have a sinister alter ego. The latest findings suggest that the protein aids tumour growth.

Susan Lindquist of the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts, and her colleagues studied mice that lack the *Hsf1* gene. These mice developed fewer tumours when they had cancer-linked mutations or when they were exposed to carcinogens than did normal mice. HSF1 also aided human tumour cell growth in culture.

The results have mixed therapeutic implications. They suggest that HSF1-inhibiting drugs could combat cancer. But HSF1-stimulating therapies are being explored as treatments for disorders including Parkinson's disease. Researchers will need to explore how each approach affects the problem targeted by the other.

CELL BIOLOGY

Quick release

Nature Struct. Mol. Biol. doi:10.1038/nsmb1305 (2007)

Test-tube models of a process that enables one nerve to transmit a signal to another do not accurately mimic what happens at living synapses, a study shows.

Nerves transmit impulses to each other thanks to small packets of neurotransmitters, which wait just inside the nerve cell until an influx of calcium ions triggers them to fuse with the cell membrane and release their contents into the synapse. Reinhard Jahn and his colleagues at the Max Planck Institute for Biophysical Chemistry in Göttingen

studied how synaptotagmin, the vesicle protein that detects calcium ions, accelerates vesicle fusion. As well as resolving some questions about the mechanism, they found that experiments with artificial vesicles do not mimic synaptotagmin's normal function owing to incorrectly structured complexes of proteins called SNAREs.

EVOLUTIONARY BIOLOGY

Protective custody

Proc. R. Soc. B doi:10.1098/rspb.2007.1028 (2007)

The human menopause might have evolved to protect offspring from being orphaned by the death of an elderly mother or to allow children to benefit from the care of grandmothers — or to do a bit of both. A test of these theories, using population data from two villages in the

Gambia between 1950 and 1975, provides evidence that grandmothers might indeed have evolutionary importance.

The Gambian data gave Daryl Shanley, from the Institute for Ageing and Health at Newcastle University, UK, and his colleagues a rare chance to study data from a natural cycle of birth and death, because the villages' populations had little access to medical care at that time. By adjusting the age of menopause in a population model based on these data, the team showed that having a post-menopausal grandmother improved survival of the offspring.

CHEMISTRY

Handle with care

J. Phys. Chem. B **111**, 10897–10904 (2007)

RDX is not a compound many would want to study. A powerful explosive used since the Second World War and favoured by several terrorist groups, RDX is apt to be detonated by shock. But why? Zbigniew Dreger and his colleagues at Washington State University in Pullman tried to find out. Their spectroscopic measurements of single RDX crystals subjected to shock waves show that the compression caused by the shock changes the crystal structure. They think this switch might predispose the material to chemical reactions, thereby initiating explosive decomposition.

PHYSICS

Chill out

Phys. Rev. Lett. (in the press)

Physicists hoping to see quantum mechanics affecting the motion of macroscopic objects have a new trick to try: a cooling technique reported by Kenton Brown and his colleagues



D. SHANLEY

at the National Institute of Standards and Technology in Boulder, Colorado.

Rather than using light, as in standard laser cooling methods, the researchers damp a small cantilever's vibrations through the electrical force between it and a nearby plate, which is connected to an electrical circuit that oscillates at radio frequencies. They reduced the cantilever's temperature from room temperature to 45 kelvin, but say that the method could in principle cool the device enough for quantum effects to show. This is predicted to happen when the object's thermal jitters are small. The technique might also improve the sensitivity of mechanical vibration and position sensors.

GENETICS

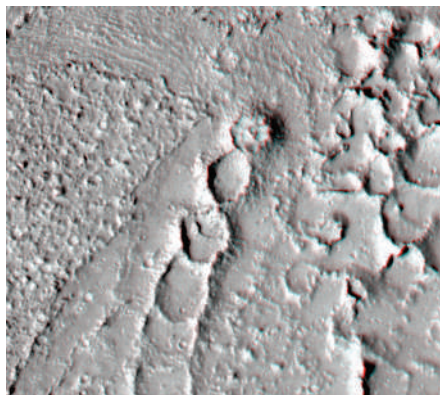
Dangerous repeats

Hum. Mol. Genet. **16**, 2326–2332 (2007)

A neurodegenerative disease called Fragile X-associated tremor and ataxia syndrome may be treatable by targeting abnormal RNA, research suggests.

The syndrome occurs in some people who have 60–200 copies of the nucleotide sequence CGG, which are transcribed into an RNA sequence that is thought to be toxic. Another disorder, called FRAXE, is caused by more than 200 repeats of the complementary sequence CCG.

Juan Botas and David Nelson of Baylor College of Medicine in Houston, Texas, and their colleagues, showed that expression of just one 90-copy CCG repeat in *Drosophila* eyes was enough to cause ocular deformities. However, flies that expressed both this



sequence and a 90-copy CGG repeat had normal eyes. The researchers propose that the two RNA transcripts bind to each other, prompting their degradation. The effect depends on a protein, Argonaute-2, that is required for the process of RNA interference.

PLANETARY SCIENCE

Rivers of rock

Science **317**, 1709–1711 (2007)

The fluid that flowed most recently through Athabasca Valles, the youngest of the 'outflow channels' on the surface of Mars, was lava and not water, according to data from the Mars Reconnaissance Orbiter.

Windy Jaeger of the US Geological Survey, Flagstaff, Arizona, and her colleagues have identified features in photos taken by the orbiter's unprecedentedly acute HiRISE camera that they can only explain in terms of volcanic flow. The researchers also say that surface features interpreted in 2005 as dust-

covered pack ice on a frozen sea are instead a pooled portion of this lava flow.

Those eager for evidence of water on Mars can take comfort in the fact that the distinctive ring-mound landforms (pictured left) spotted in these images were produced by outbursts of gas from below the lava flows. They presumably mark spots where ice or ground water in the rocks over which the lava flowed vaporized explosively.

MARINE BIOLOGY

Seaweed searches

Proc. Natl Acad. Sci. USA doi:10.1073/pnas.0704778104 (2007)

The lush kelp forests that characterize some cold ocean waters have now been discovered in tropical locales. The finding overturns the prevailing hypothesis that tropical waters are too warm and nutrient deficient to support the large seaweeds.

Michael Graham of Moss Landing Marine Laboratories in California, and his colleagues, used oceanographic data to model kelp habitats. The model accurately located known kelp forests, but also predicted more than 23,500 square kilometres of potential kelp habitat in deep tropical regions where upwelling brings in nutrients and clear waters allow sufficient light to penetrate. The researchers then searched the waters near the tropical Galapagos Archipelago and found several new kelp forests, as predicted.

These tropical kelp forests could provide unique biodiversity hotspots that may warrant inclusion in marine conservation programmes.

JOURNAL CLUB

David K. Campbell
Boston University, USA

A physicist highlights a three-in-one deal for nonlinear science

As a student of nonlinear phenomena, I am continually amazed by new examples of deterministic chaos, solitary waves and fractals.

A recent study (R. H. Goodman and R. Haberman *Phys. Rev. Lett.* **98**, 104103; 2007) gave me the rare pleasure of seeing all three of these fundamental manifestations of nonlinearity woven together.

This paper addresses the collisions of solitary waves — localized nonlinear waves that

propagate without changing shape and are found in systems ranging from solids to optical fibres.

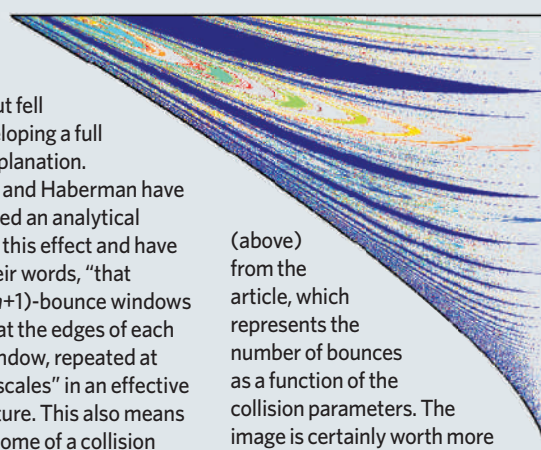
In the 1980s, with several colleagues, I studied this problem numerically (see, for example, D. K. Campbell and M. Peyrard *Physica D* **18**, 47–53; 1986). We discovered a surprising 'bounce' phenomenon, in which solitary waves would collide, remain trapped for a number (n) of bounces and then escape to infinity. This behaviour occurred only when the waves had specific relative velocities on colliding; these bounce windows were interspersed with regions in which the waves repelled each other immediately.

We developed a heuristic explanation for this behaviour, consistent with the waves behaving

like elastic particles that can be deformed, but fell short of developing a full analytical explanation.

Goodman and Haberman have now developed an analytical treatment of this effect and have shown, in their words, "that clusters of $(n+1)$ -bounce windows accumulate at the edges of each n -bounce window, repeated at diminishing scales" in an effective fractal structure. This also means that the outcome of a collision is exquisitely sensitive to the initial velocity, a hallmark of deterministic chaos.

If all the above seems dry, take a look at the wonderful graphic



(above) from the article, which represents the number of bounces as a function of the collision parameters. The image is certainly worth more than these few hundred words.

Discuss these papers at <http://blogs.nature.com/nature/journalclub>

NEWS

Chemists poke holes in ozone theory

As the world marks 20 years since the introduction of the Montreal Protocol to protect the ozone layer, *Nature* has learned of experimental data that threaten to shatter established theories of ozone chemistry. If the data are right, scientists will have to rethink their understanding of how ozone holes are formed and how that relates to climate change.

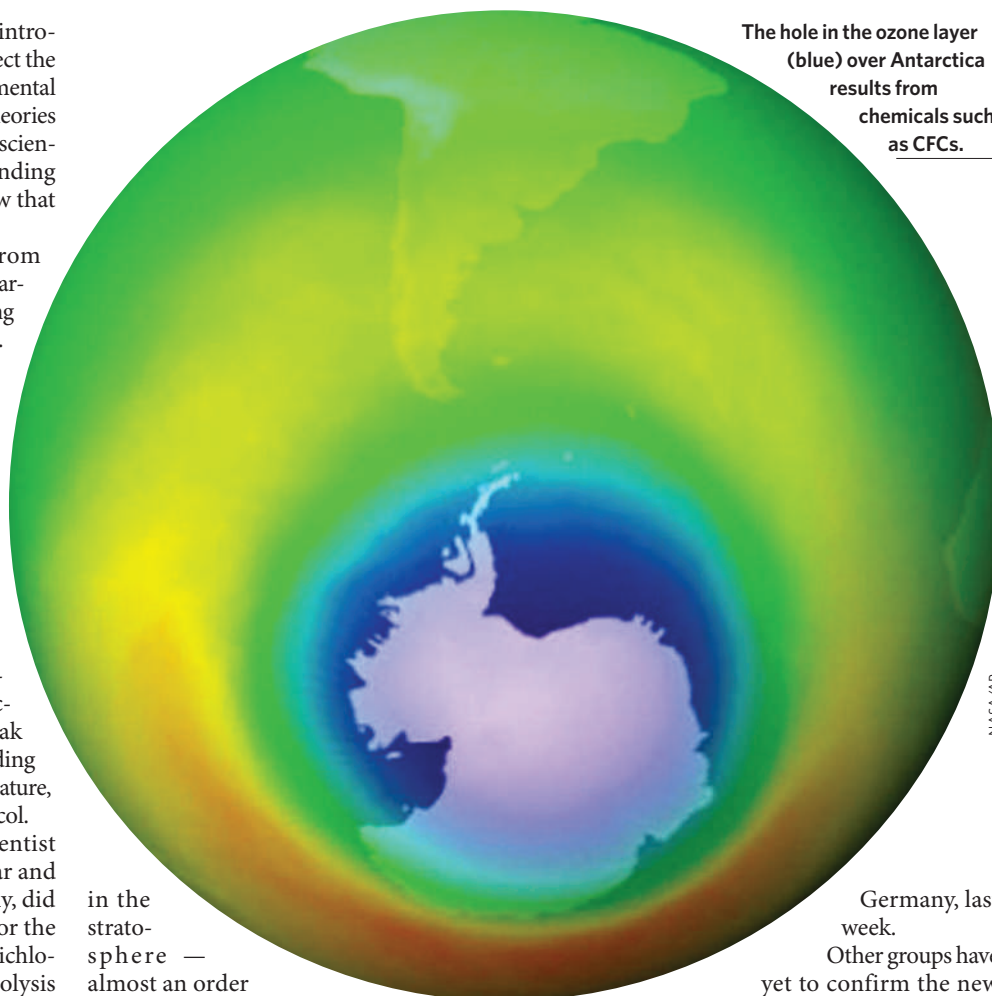
Long-lived chloride compounds from anthropogenic emissions of chlorofluorocarbons (CFCs) are the main cause of worrying seasonal ozone losses in both hemispheres. In 1985, researchers discovered a hole in the ozone layer above the Antarctic, after atmospheric chloride levels built up. The Montreal Protocol, agreed in 1987 and ratified two years later, stopped the production and consumption of most ozone-destroying chemicals. But many will linger on in the atmosphere for decades to come. How and on what timescales they will break down depend on the molecules' ultraviolet absorption spectrum (the wavelength of light a molecule can absorb), as the energy for the process comes from sunlight. Molecules break down and react at different speeds according to the wavelength available and the temperature, both of which are factored into the protocol.

So Markus Rex, an atmosphere scientist at the Alfred Wegener Institute of Polar and Marine Research in Potsdam, Germany, did a double-take when he saw new data for the break-down rate of a crucial molecule, dichlorine peroxide (Cl_2O_2). The rate of photolysis (light-activated splitting) of this molecule reported by chemists at NASA's Jet Propulsion Laboratory in Pasadena, California¹, was extremely low in the wavelengths available

in the stratosphere — almost an order of magnitude lower than the currently accepted rate. "This must have far-reaching consequences," Rex says. "If the measurements are correct we can basically no longer say we understand how ozone holes come into being." What effect the results have on projections of the speed or extent of ozone depletion remains unclear.

The rapid photolysis of Cl_2O_2 is a key reaction in the chemical model of ozone destruction developed 20 years ago² (see graphic). If the rate is substantially lower than previously thought, then it would not be possible to create enough aggressive chlorine radicals to explain the observed ozone losses at high latitudes, says Rex. The extent of the discrepancy became apparent only when he incorporated the new photolysis rate into a chemical model of ozone depletion. The result was a shock: at least 60% of ozone destruction at the poles seems to be due to an unknown mechanism, Rex told a meeting of stratosphere researchers in Bremen,

The hole in the ozone layer (blue) over Antarctica results from chemicals such as CFCs.

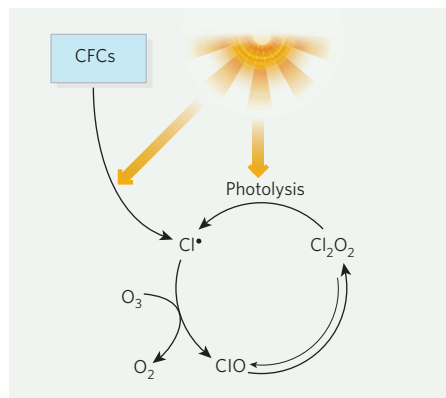


Germany, last week.

Other groups have yet to confirm the new photolysis rate, but the conundrum is already causing much debate and uncertainty in the ozone research community. "Our understanding of chloride chemistry has really been blown apart," says John Crowley, an ozone researcher at the Max Planck Institute of Chemistry in Mainz, Germany.

"Until recently everything looked like it fitted nicely," agrees Neil Harris, an atmosphere scientist who heads the European Ozone Research Coordinating Unit at the University of Cambridge, UK. "Now suddenly it's like a plank has been pulled out of a bridge."

The measurements at the Jet Propulsion Laboratory were overseen by Stanley Sander, a chemist who chairs a NASA panel for data evaluation. Every couple of years, the panel recommends chemical kinetics and photochemical data for use in atmosphere studies. Until the revised photolysis rate has been evaluated, which won't be before the end of next year, "modellers must make up their minds about what to do,"



Cl_2O_2 is key to ozone (O_3) depleting reactions such as this one, in which photolysis results in a chlorine radical (Cl^*) that reacts with O_3 .

says Sander. One of the problems with checking the data is that the absorption spectra of chloride compounds are technically challenging to determine. Sander's group used a new technique to synthesize and purify Cl_2O_2 . To avoid impurities and exclude secondary reactions, the team trapped the molecule at low temperatures, then slowly warmed it up.

"Reactions in experimental chambers are one thing — the free atmosphere is something else," says Joe Farman, one of the scientists who first quantified the ozone hole over Antarctica³. "There's no doubt that ozone disappears at up to 3% a day — whether or not we completely understand the chemistry." But he adds that insufficient control of substances such as halon 1301, used as a flame suppressor, and HCFC22, a refrigerant, is a bigger threat to the success of the Montreal Protocol than are models that don't match the observed losses.

Hot topic

Meanwhile, atmosphere researchers have started to think about how to reconcile observations of ozone depletion with the new chemical models. Several thermal reactions, or combinations of reactions, could fill the gap. Sander's group has started to study possible candidates one by one — but so far without success.

Rex thinks that a chemical pathway involving a Cl_2O_2 isomer — a molecule with the same atoms but a different structure — might be at play. But even if the basic chemical model of ozone destruction is upheld, the temperature dependency of key reactions in the process could be very different — or even opposite — from thought. This could have dramatic consequences for the understanding of links between climate change and ozone loss, Rex says.

The new measurements raise "intriguing questions", but don't compromise the Montreal Protocol as such, says John Pyle, an atmosphere researcher at the University of Cambridge. "We're starting to see the benefits of the protocol, but we need to keep the pressure on." He says that he finds it "extremely hard to believe" that an unknown mechanism accounts for the bulk of observed ozone losses.

Nothing currently suggests that the role of CFCs must be called into question, Rex stresses. "Overwhelming evidence still suggests that anthropogenic emissions of CFCs and halons are the reason for the ozone loss. But we would be on much firmer ground if we could write down the correct chemical reactions."

Quirin Schiermeier

1. Pope, F. D., Hansen, J. C., Bayes, K. D., Friedl, R. R. & Sander, S. P. *J. Phys. Chem. A* **111**, 4322–4332 (2007).
2. Molina, L. T. & Molina, M. J. *J. Phys. Chem.* **91**, 433–436 (1987).
3. Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).

Pressure for environmental disclosure increases

Companies are increasingly acknowledging the risks posed by global warming, suggest surveys released on 24 September. The latest Global Corporate Climate Change Report, released by the advocacy group Carbon Disclosure Project (CDP) at an event in New York headlined by former US President Bill Clinton, details disclosures of energy costs from more than 1,300 companies around the world (see graphic).

The London-based group of institutional investors is campaigning for the right of shareholders to probe the environmental policies of firms in which they invest. It argues that the possibility of real damages from global warming and the almost certain introduction of regulations penalizing emissions of carbon dioxide and other greenhouse gases mean that climate now poses a material liability to many companies that should

be publicly disclosed.

CDP is leveraging investor pressure to push for voluntary disclosure, but others have opted for brute force. New York Attorney General Andrew Cuomo issued subpoenas on 14 September for internal documents from five companies planning to build new coal-fired power plants, questioning whether they have fully disclosed the financial risks of increasing their greenhouse-gas emissions. Ceres, a coalition of environmentalists and investors, and Environmental Defense, a non-profit organization based in New York, followed up days later with a petition to the US Securities and Exchange Commission seeking to increase financial disclosure requirements on corporate exposure to 'climate risk'.

On Monday, Wal-Mart announced that it intends to ask its suppliers for the

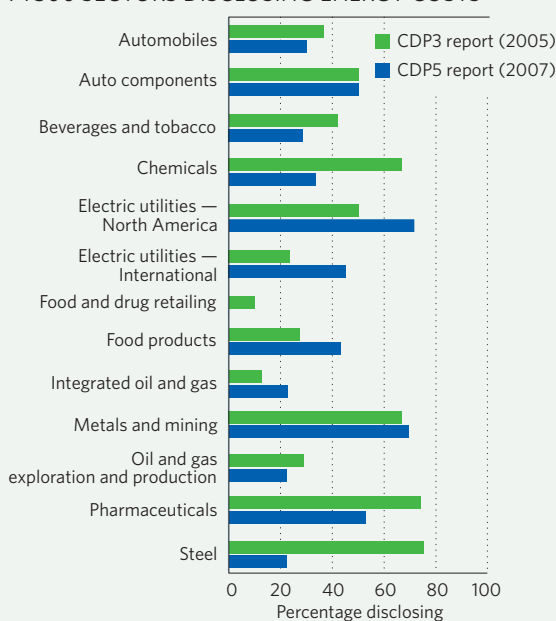
energy costs involved in manufacturing seven of its product lines. The retail chain will then work with CDP to translate this information into greenhouse-gas emission figures.

"It will be money coordinated through Wall Street that builds the energy-efficient economy of the twenty-first century," says CDP head Paul Dickinson. "There are risks, but there are also big opportunities."

CDP asked firms to assess the risks — and opportunities — presented by global warming, and to describe any strategies being used to reduce greenhouse-gas emissions. The survey, signed by more than 300 institutional investors managing some US\$41 trillion in assets, found that 76% of respondents reported implementation of a programme to reduce greenhouse-gas emissions, compared with 48% a year ago. The percentages of respondents who thought global warming posed an economic threat and of those who believed it offered opportunities were roughly equal, at 79% and 82%, respectively.

This is the fifth annual report by CDP, and the statistics indicate some success: 77% of the world's 500 largest companies responded, compared with 47% in 2002. European firms led the way with 83% participating. North American firms gained some ground with a 74% response rate, compared with 66% last year. Among the large firms that did not respond were Apple Computer, Philips Electronics and Gazprom. Jeff Tollefson

FT500 SECTORS DISCLOSING ENERGY COSTS



SCORECARD

**Emoticons**

These infuriating yet useful keystroke combinations have reached their silver jubilee. They were first used in 1982 by Carnegie Mellon University's Scott Fahlman, who proposed them as an aid to electronic communication :-)

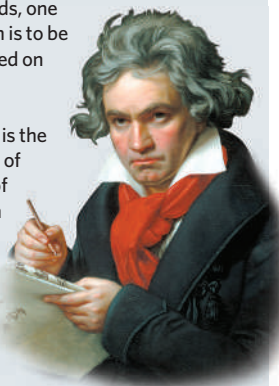
**Conventional languages**

Meanwhile, 20% of the world's 7,000 languages are in danger of disappearing, with one going extinct every 14 days on average :-)

NUMBER CRUNCH

10 is the number of strands of Ludwig van Beethoven's hair that have been used to create three diamonds, one of which is to be auctioned on eBay.

0.56 is the number of carats of the gem on sale.



\$500,000 is the amount it is expected to fetch — meaning that you'll need a modern rock star's fortune if you want to be the owner.

ON THE RECORD

“The International Space Station is an orbital turkey.”

Nobel-prizewinning particle physicist Steven Weinberg is less than impressed with the achievements of manned spaceflight.

OVERHYPED

Velociraptor

Continuing the turkey theme, palaeontologists have announced that velociraptor — far from being fearsome or scaly — actually more closely resembled the traditional holiday poultry. Did Michael Crichton mislead us all?

Sources: Associated Press, LA Times, Daily Telegraph, Space.com, Science



E. NATHAN/ALAMY

The University of Cape Town has received the lion's share of research positions from the government.

South African scheme lures in top talent

CAPE TOWN

South Africa's mission to create a globally competitive national academy of sciences reached another milestone last week with the announcement of 51 new research chairs. The South African Research Chairs Initiative, launched by the government in December 2006, aims to create 210 research chairs by 2010 in a range of disciplines from nanophotonics to dignity jurisprudence.

The government hopes that the scheme will seduce foreign researchers to come to South Africa, as well as encourage the country's top researchers to stay and expatriates to return. Two-thirds of the 72 appointments made so far are people already working in South Africa, but the eventual aim is for 60% of the appointments to come from abroad. To redress the balance, the next round of awards is likely to be confined to applicants from outside the country, says Robin Drennan of South Africa's National Research Foundation, which administers the scheme.

But Cheryl de la Rey, deputy vice-chancellor for research at the University of Cape Town, says that the selection process makes it difficult to attract overseas candidates. “It's not easy to get people to move continents, and a problem with the current procedure is that it takes so long that some foreign-based candidates take up other offers. We must try to improve the turnaround time,” she says.

The National Research Foundation's chief executive Mzamo Mangaliso says the selection process will be overhauled to speed it up without compromising quality. “We hope to empower our universities to recruit suitable candidates in a speedy and focused manner,” he says.

Although 14 of the country's 22 universities have been awarded at least one chair, two-thirds

of the appointments are concentrated at three universities: the University of Cape Town, the University of the Witwatersrand in Johannesburg and the University of Stellenbosch. Piet Steyn, Stellenbosch's director of research, says that this is unavoidable because a critical mass of researchers is needed as well as access to infrastructure such as experimental facilities and libraries — which are concentrated within relatively few regions of the country. The chairs, which will cost the state 170 million rand (US\$24 million) over the next five years, are a very positive and important intervention, Steyn says.

So far, around two-thirds of the incumbents are white, but the initiative aims to have 60% of the places filled by black researchers. One of the advantages of appointments from within the country, says de la Rey, is that universities are obliged, by the terms of the scheme, to fill the new chair-holder's previous post. This can be done at a more junior level, where it is often easier to diversify in terms of ethnicity and gender, she says.

Cancer biologist Iqbal Parker of the University of Cape Town, who is director of the newly established International Centre for Genetic Engineering and Biotechnology there, is delighted with his new research chair. He and Frank Brombacher, another research-chair holder, will head up the centre's first two research programmes — on non-communicable diseases, which are increasing in Africa, and the immunology of infectious diseases, respectively. Parker says that the main advantage of the scheme is that the research-chair positions incorporate awards of studentships, postdoctoral fellowships and running expenses for five years.

Michael Cherry

Q&A

King of the stem cells

On 14 September, the California Institute for Regenerative Medicine (CIRM) appointed Australian biologist **Alan Trounson** as its new president. The CIRM has lacked a permanent director since neuroscientist Zach Hall departed in April, amidst rumours of tension between himself and Robert Klein, chair of the \$3-billion agency's board.

Why were you picked?

I think the CIRM board was interested in my perspective of wanting to push very hard at the pipeline between discovery and clinical application. I am more interested in getting discoveries through to the clinic than simply commercial operations or research for the sake of research.

What's exciting you in the field?

There are a lot of options that are brimming with potential. Mesenchymal [multipotent] stem cells are already in clinical trials.

Embryonic stem cells (ESCs) are coming of age and I'm interested in the clinical trials proposed by Geron, based in Menlo Park, California, and the University of California, Irvine.

Does the CIRM need to recalibrate its vision because companies don't see ESC therapies as viable?

Companies have been a bit too optimistic by looking for cures in very short timeframes. I'd rather the commercialization be a natural spin-off of the pipeline and that's what I think will happen in California.

Given the international expectations and local politics, what is possible at the CIRM?

I've got to live with the current [CIRM] vision and maximize the efficiency of delivery of this medical opportunity to patients. I've got to link in with researchers in other US states, but we should treat this as a global exercise. California will lead it, because they've got such a lot of money and researchers. But if we incorporate our efforts, and minimize duplication and maximize output, this medical revolution will happen more quickly than we expected.

What did you learn from your involvement in politics in Australia?

Be very, very careful what you say. And utilize the politically astute people around you to get the message delivered, rather

than trying to take the whole system head on. I would not be inclined to [get involved in politics] again. It was hard going and stressful and clearly I didn't get it right every time, and I apologized for that.

Do you think the CIRM's dual leadership structure could be awkward?

I've always worked in partnership with people. I worked with Carl Wood to deliver *in vitro* fertilization. I worked with David de Kretser to set up the Institute of Reproduction and Development at Monash

University. All my life I've had partnerships with people who bring a different set of skills from my own. Bob Klein is incredibly good at what he does. He's persuasive, he has a tremendous memory and he's got political connections. I don't have any of that, so creating a partnership with Bob is exactly what should be done.

Is giving up the lab difficult?

That's hard, but it's time. I could get another paper in *Nature* or somewhere else, and that's a thrill. But I've made my contribution in the lab. To have a chance to contribute at this level is very exciting.

Why do you call this the best job in stem cells in the world?

You're working with the best researchers, you've got an enormous amount of money and you've got political will and the people of California behind you. It's a huge challenge, but I think the resources are there to accommodate it.

When will be the first stem-cell therapy?

I see it as a continuum. Adult stem cells are happening. Embryonic stem cells will come into use, and they won't be immediate cures for everything. You need drugs and protocols as well as the cells, and you've got to work with the immune system. ■

Interview by Erika Check Hayden.



Alan Trounson wants stem-cell therapy clinically applied.

BLOOMBERG NEWS/LANDOV/PHOTOSHOT



WRIST BONES BOLSTER HOBBIT STATUS

Ape-like wrists suggest that *Homo floresiensis* was a distinct species.

www.nature.com/news

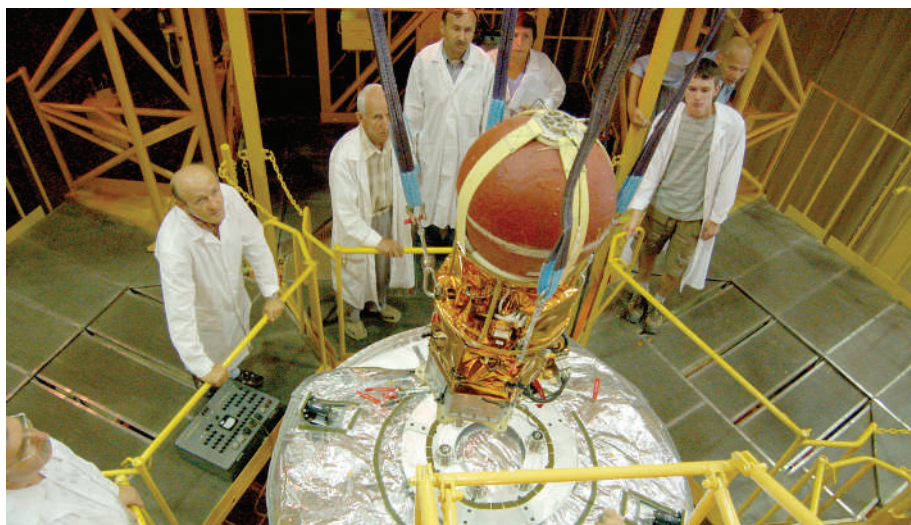
SCIENCE

Dropping a line from space

As *Nature* went to press, scientists were scrambling to locate a small capsule from space now believed to be somewhere in Kazakhstan. The capsule is part of an ambitious experimental space-mail delivery system that aimed to use a 30-kilometre-long satellite tether — the longest yet. In the early hours of 25 September, the ribbon from the satellite orbiting 300 kilometres above Earth was cut and the capsule parachuted to Earth, although off-target.

The test forms part of the European Space Agency's Foton-M3 mission, which set out on its 12-day Earth-orbiting science mission on 14 September. Foton-M3 is carrying more than 40 science experiments, including the tether test, called Young Engineers' Satellite 2 (YES2), a €2.7-million (US\$3.8-million) project involving more than 450 students. The idea was to demonstrate the feasibility of a cheap way of returning small objects to Earth from distances similar to that of the International Space Station some 330 kilometres away.

Delivering small payloads from low Earth orbits is not as easy as simply dropping them. "You don't want this object to spin around the Earth," says Francesco Emma, head of the European Space Agency's education office. The capsule, or any other payload, must slow down so that it can come out of its orbit. This can be done using a retro-rocket, but tethers are another way. "You can lower the capsule," says Emma. "When it is released it doesn't have enough energy to spin."



YES2, a 6-kilogram capsule (pictured), was hung from a satellite by a tether before falling to Earth.

The tether, just 0.6 millimetres thick, was made from a super-strong polyethylene fibre called Dyneema and weighed 5.5 kilograms. YES2 consisted of three parts: the 6-kilogram capsule, together with heat shield, that returned to Earth; a carrier that held the capsule in place; and a deployer, bolted to Foton-M3, that held both capsule and carrier and a spool of tether.

The capsule was deployed in two stages — 3.5 kilometres of tether was released at first. After the landing location was targeted, the idea was to reel out the rest of the tether. Once the

pendulum-like tether and capsule had swung to the vertical, the tether would be cut, and the capsule would parachute to Earth in about 30 minutes. But the rate of cable uncoiling was 5 metres per second, not 12 metres per second as planned, so the cable unwound only 8.5 kilometres, Russian reports suggest — although this was still the first successful tether mission to Earth. A transmitter on the satellite should allow mission controllers to locate the capsule's landing site.

Katharine Sanderson

Stem-cell fraudster 'is working in Thailand'

Disgraced South Korean cloning scientist Woo Suk Hwang has set up a research base in Thailand, according to reports from South Korea last week. Science-policy officials in Thailand say they are worried about the reports, as they knew nothing about this.

Se Pil Park, a fertility expert at Cheju National University in South Korea, says Hwang and ten colleagues have gone to Thailand to carry out cloning experiments involving the transfer of human DNA into eggs from cows, pigs and other animals. The resulting embryos cannot produce viable offspring, but might result in embryonic stem-cell lines that could be cultured for



Woo Suk Hwang is said to have moved to Thailand to continue cloning.

research and for therapeutic use.

In 2004, Hwang claimed to have produced embryonic stem cells by cloning human DNA using human eggs. But the data behind these claims were found to have been

fabricated. Park would not disclose details of the researcher's latest work, nor where he was based.

Hwang, who remains on trial for fraud, embezzlement and violation of a bioethics law related to his discredited experiments, lost his licence to do research with human eggs. According to Park, Hwang went to Thailand to carry out interspecies cloning because he thought he would face opposition from ethicists in his home country.

In Thailand, human reproductive cloning was made illegal in 2002, but there are currently no laws regulating therapeutic cloning involving human DNA or interspecies cloning. According to Thai veterinary

scientist Pradon Chatikavanij, proposed regulations on animal experimentation now before the main legislative body, the Council of State, might restrict interspecies cloning, although no decision on this has yet been made. Chatikavanij is the former head of and now a consultant for the National Research Council of Thailand, which is in charge of drafting the regulations.

Chatikavanij says the Thai government knows nothing about any research being done by Hwang in Thailand. "We should know what they are doing. We are very concerned about what is happening," he says.

David Cyranoski

YONHAP, S. R. BAK/AP

Bacteria may be wiring up the soil

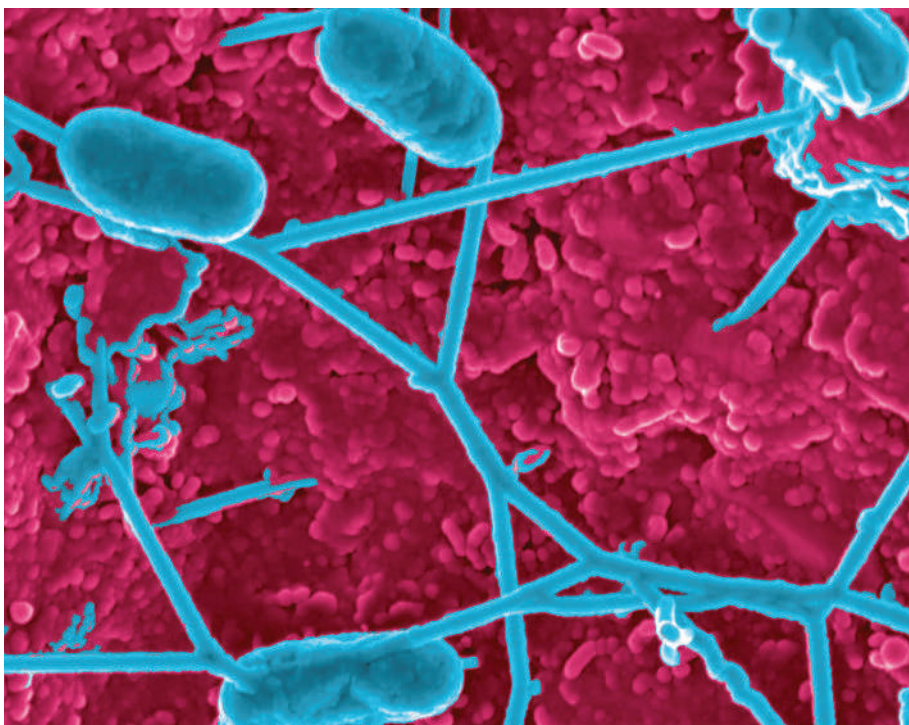
Bacteria can sprout webs of electrical wiring that transform the soil into a geological battery, a team of researchers claims. Some soil bacteria form networks of tiny wires linking individual bacterial cells into a web-like electrical circuit, they report (D. Ntarlagiannis *et al. Geophys. Res. Lett.* **34**, L17305; 2007). The wires allow the bacteria to get rid of electrons generated during metabolism, transporting them to distant 'electron dumps'.

The earth beneath our feet might act as a gigantic circuit built by microbes to power their metabolic systems, suggests Yuri Gorby, a geochemist previously at the Pacific Northwest National Laboratory in Richland, Washington, and now at the J. Craig Venter Institute in La Jolla, California. As the bacteria consume nutrients for energy, they spit out unwanted electrons into the circuit.

Last year, Gorby and his colleagues discovered that *Shewanella oneidensis* bacteria can grow long filaments, just 100 nanometres (a hundred millionths of a millimetre) thick, which conduct electricity (Y. A. Gorby *et al. Proc. Natl Acad. Sci. USA* **103**, 11358–11363; 2006). The researchers presented evidence that the microbes use these 'nanowires' to shunt electrons produced during metabolic reactions onto the surface of mineral grains in the soil, to be taken up by metal ions. Without an electron acceptor, the bacteria cannot function properly and die. The researchers found that several other bacterial species also produce such nanowires.

Oxygen molecules act as convenient electron dumps for bacteria that lie near the soil surface. But little air penetrates to some environments, such as deep lake sediments or waterlogged soils. Now, Gorby and his team think they have found evidence that the bacterial nanowires can link up into a network, conducting electrons to the aerated surface. The researchers filled plastic columns with wet sand infiltrated with a nutrient compound (lactate), and allowed *S. oneidensis* to grow in this 'fake soil'. Only the top of the column was in contact with air.

Electrodes inserted at various heights up the columns revealed that, after about ten days, electrical charge was coursing up the column. Gorby's team examined the sand under a microscope and found that it was threaded by a web of filaments between the bacterial cells. These are wires that provide the pathways for electron transport up to the surface, they suggest.



Connected: false-colour image showing bacterial cells apparently linked by filamentary 'nanowires'.

In contrast, when the team grew a colony of mutant cells that could spawn only very thin, frail and non-conducting filaments, the electrodes in the column remained uncharged.

"This is a new aspect of microbiology," says Gorby. "I think we will find that it's a dominant lifestyle for microbes to live in an electrically connected community." Sediment batteries, which power low-energy devices such as sensors, might unwittingly be exploiting these electric webs, he says. In these devices, electrodes are inserted into underwater sediments to pick up tiny currents generated by microbial activity. "Our work shows how these electrons are being mobilized," Gorby says.

"If this idea is right, it is really quite remarkable," says Kenneth Nealson, a geobiologist at the University of Southern California in Los Angeles, who collaborated with Gorby on last year's discovery of microbial nanowires.

But not everyone is persuaded by the claim. "I think many of us would like to believe that microbial communities may be hard-wired, because it is just such an attractive and fantastic notion," says Derek Lovley, a microbiologist at the University of Massachusetts at Amherst.

"But there is just no evidence for it." He adds that most microbes in anaerobic environments don't need long-distance access to oxygen in order to get rid of the electrons generated when they break down organic matter.

Besides, he says, the experiments don't show conclusively that the filaments are acting as wires, nor that this is what allows electrons to move through the system. He thinks the microbes probably shed electrons by releasing soluble electron shuttle molecules, not by passing them along nanowires.

"It is well documented that *Shewanella* releases soluble molecules that act as electron shuttles," he says. "The difference between the wild-type and the mutant could be attributed to the mutants' inability to react with a soluble electron shuttle. This is a more plausible explanation."

Bruce Logan, a microbiologist at Pennsylvania State University in University Park, who has worked on microbe-driven sediment batteries, shares such reservations. "I believe they see nanowires, but I see no evidence that a long-distance connection is established here," he says. "It may be possible, but these data don't prove it."

"There are clearly way more questions raised than answered at this point," Nealson admits. ■

Philip Ball

B. AREY & E. HILL/PACIFIC NORTHWEST NAT'L LAB.



NATURE REPORTS STEM CELLS

The news behind the science,
the science behind the news.
[www.nature.com/
stemcells](http://www.nature.com/stemcells)

S. GSCHMEISSNER/SPL

Stem cells by any other name

The tricky word 'embryonic' is to be removed from the name of the US Human Embryonic Stem Cell Registry, the National Institutes of Health announced last week. It will be renamed the Human Pluripotent Stem Cell Registry.

The change results from an executive order issued by President George W. Bush after he vetoed legislation that would have lifted some restrictions on federal funding for human embryonic stem-cell research (see *Nature* doi:10.1038/news070618-13; 2007). The order was widely viewed as an attempt to downplay the potential of human embryonic stem cells. Even its title is politically charged: "Expanding Approved Stem Cell Lines in Ethically Responsible Ways."

But Charis Thompson, a professor of rhetoric and a member of the Berkeley Stem Cell Center in California, says both sides play at that game: advocates of embryonic stem-cell research avoid using terms such as

'embryo' and 'cloning', she observes. "This kind of linguistic deflation of public anxiety is a hallmark of moving ahead in this field in the United States and will only intensify," Thompson says.

Besides removing 'embryonic' from the names of the cell lines that can receive federal research funding — as instructed in the executive order — the 18 September announcement sets out a strategy of awarding grants to create cell lines without the involvement of embryos, adding more funds to existing grants and "aggressively pursuing an assessment of the potential of alternative sources of pluripotent stem-cell lines".

This rankles with many scientists working to create pluripotent stem-cell lines, who say that such an assessment is not possible without research on newly derived embryonic stem-cell lines. "It's not

an alternative; these are complementary technologies," says Rudolf Jaenisch at the Whitehead Institute in Cambridge, Massachusetts, who is working to reprogramme non-embryonic cells to be as flexible as embryonic stem cells.

One problem is that pluripotency is a notoriously imprecise term. Not even human embryonic stem cells have been established as pluripotent, as researchers have yet to establish their ability to become all human cell types, says

Glyn Stacey, head of the UK Stem Cell Bank in Hertfordshire, which handles stem-cell lines created from embryos as well as those derived from fetal and adult tissues.

The first non-embryonic stem-cell lines could be added to the registry in the next few months.

Monya Baker

See editorial, page 377.

**"Pluripotency
is a notoriously
imprecise term."**

Clinical Pharmacology & Therapeutics

Read others' thoughts. **Discuss** points of view. **Connect** in clinical pharmacology

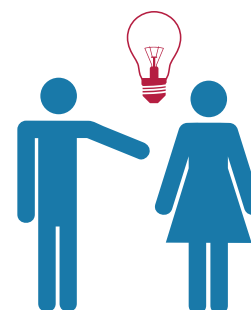
Why use Nature Network?

- **Create** a personal profile page for you and your research
- **Build** your own global online network
- **Form** a group for your lab, department/division, institution or organization
- **Search** comprehensive listings of upcoming seminars and conferences
- **Post** your own events to local calendars

Clinical Pharmacology & Therapeutics, now on Nature Network.

Extend your thoughts and share your knowledge.

<http://network.nature.com/group/cpt>



nature publishing group 

Prospects power up for nuclear energy

American energy utility NRG Energy is planning to build two reactors for nuclear power in Texas. Its application to the US Nuclear Regulatory Commission for permission to construct the reactors marks the first such request in the United States for three decades.

Meanwhile, Italy is reversing a 20-year post-Chernobyl moratorium on nuclear fission research to participate in the scientific development of 'Generation IV' nuclear reactors. How exactly the country will contribute to the Generation IV research has not yet been revealed. The planned next generation of nuclear power plants is intended to improve nuclear safety and to minimize waste and natural-resource use compared with the current generation of plants.

Europe plots course for funding navigation system

The European Commission is launching a bid to save Galileo, the continent's troubled satellite navigation system. It has asked the European Parliament to back its plan to foot the whole bill for the 30-satellite system. The cost would be some €3.4 billion (US\$4.8 billion) over the next six years.

Galileo would see Europe break free from its dependence on the US-based Global Positioning System (GPS). But the project suffered a serious setback earlier this year when a public-private deal to fund the system collapsed (see *Nature* 447, 765; 2007).

If the commission manages to salvage Galileo, it would get the extra cash from surplus monies in the agricultural and administration budgets. Member states will take the first steps to making a decision on the proposals in early October.

Meanwhile, President George W. Bush has announced that the United States will permanently leave the capability

Meteorite proves to be a hit in Peru

Researchers are examining a 17-metre-wide crater in southern Peru, looking for fragments of what is believed to be one of the largest meteorites to hit land in years.

The meteorite apparently plummeted to Earth on 15 September near the village of Carancas, west of Lake Titicaca on the Bolivian border. Scientists from Peru's Geological, Mining and Metallurgical Institute in Lima last week checked the crater (pictured), now filled with water, for debris from the object they described as a chondrite or stony meteorite.

People living near the impact complained of odd smells and illnesses, but some scientists speculated that the reported sickness may be group hysteria. They noted that objects that come from the sky have deep cultural significance to villagers.



LA REPUBLICA/AP

of distorting signals off any future GPS satellites. This feature was designed to give the US government a better working version of GPS than civilians, but has not been used since 2000.

HIV vaccine failure prompts Merck to halt trial

An HIV vaccine being developed by Merck has apparently failed, causing the company to halt a large and once-promising clinical trial last week.

Merck's STEP vaccine used a mixture of components from three weakened adenoviruses to carry three synthetically produced HIV genes. The hope was that each gene would stimulate an immune response against the virus, as earlier trials had suggested.

The latest trial began in 2004 and enrolled 3,000 people considered to be at high risk of infection. But a group of 741 volunteers who received the vaccine saw 24 HIV infections, compared with the control group of 762 people who saw 21 infections. Furthermore, the vaccine did not reduce the amount of HIV in the bloodstream of those infected.

FDA poised for broader powers over drugs on sale

The US Congress last week passed a bill significantly increasing the powers of the Food and Drug Administration (FDA) to police the safety of marketed drugs (see *Nature* 446, 844–845; 2007). President George W. Bush is expected to sign the bill by 30 September, when a current law expires.

Under the new law, the FDA will for the first time be able to order companies to make label changes on marketed drugs; to run new clinical trials when

safety issues emerge; and to fine non-compliant companies up to \$10 million. It also includes \$225 million in new drug-company-paid user fees that the agency will spend on post-market drug safety.

The bill requires the FDA to cut by 25% the number of waivers it grants to external experts with financial interests in matters on which they are advising the agency. Such waivers allow them to serve on committees that advise the agency on drug approvals, withdrawals and label changes — advice it almost always follows.

California universities maintain tobacco habit

Under a policy adopted last week, scientists at the University of California's ten campuses can continue to accept research grants from tobacco firms. But new grants coming from the industry will undergo added scrutiny.

The University of California's governing board approved the policy on 20 September after several years of intense debate — spurred by a core of outraged scientists (see *Nature* 446, 242; 2007). Under the new policy, proposed tobacco-firm grants will undergo peer review by panels of scholars at each campus, where the chancellors will oversee the monitoring process. There will also be an annual report on all research funded by the tobacco industry, describing both submitted and approved projects.

Last year, seven of the university's campuses received a total of \$16 million for 23 projects, all funded by Philip Morris of Richmond, Virginia.

Correction

The Editorial 'Epo, by any other name' (*Nature* 449, 259; 2007) incorrectly stated that the drug Cerezyme is produced by Genentech. It is, in fact, produced by Genzyme Therapeutics.



Up in the air: Europe's planned Galileo navigation system has yet to secure full funding.

Passing the test

What role should the federal government have in pre-college science education? **David Goldston** looks at why the US Congress is acting now to help define that.

On 9 August, President George W. Bush rather grudgingly signed into law the America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education and Science Act. The scientific community has tended to focus on provisions that authorize doubling the budgets for the National Science Foundation (NSF), the Department of Energy's (DoE's) science programmes and the National Institute of Standards and Technology. But most of the statute's 148 pages are devoted to creating new and expanded programmes in pre-college science and mathematics education.

The focus on education should be welcomed and cannot be taken for granted. Congressional enthusiasm for pre-college science education has waxed and waned over the years (see *Nature* **446**, 714; 2007). A high point was the National Defense Education Act, passed in 1958 in the wake of the Soviet Sputnik satellite mission, which inaugurated federal funding for the topic.

But in the early years of the Reagan administration, Congress eliminated the NSF's pre-college education programmes in an effort to stem controversy over curricula and to rein in federal spending. That decision was reversed within a couple of years, but the programmes took some time to recover. About a decade later, Congress instructed the DoE to close down its pre-college education programmes, saying that they detracted from the agency's primary missions. In 2001, science education seemed to be an after-thought when Congress passed the No Child Left Behind Act, which requires testing pupils to evaluate the quality of schools. And as recently as 2005, Congress set up an inter-agency panel with an eye to pruning federal programmes in the area, implying that there were too many and that they were inadequately coordinated and evaluated.

So why has Congress now focused on science education? The most immediate cause of the turnaround was the Democratic takeover of Congress this year. Although plenty of Republicans in previous Congresses supported a revived federal role in pre-college science, the House Republican leadership blocked consideration of a competitiveness bill because vocal congressional conservatives and the White House opposed spending more and creating new programmes. (The main power of party leadership



PARTY OF ONE

is its authority to schedule what comes before Congress for a vote.) But the longer-term change was aided by the willingness of business groups such as the Business Roundtable, which comprises the chief executives of around 160 major companies, to put some of their clout behind lobbying for improving pre-college education.

Now the question is whether enough has changed to make a difference. A law designed to bolster the NSF's role in education was enacted in 2002, but funding for the programmes didn't follow. The good news this time is that Congress seems to be willing to put its money where its mouth is. Next year's spending bills, for example, would provide the NSF's education programmes with their first real increase in four years — according to the American Association for the Advancement of Science. The Senate has the higher figure — \$850 million for the fiscal year that starts 1 October. And the programmes would increase at a faster rate than the agency as a whole. (The president is threatening to veto most of the upcoming spending bills, claiming that they are too large, so it's much too early to guess what will actually be signed into law.)

But not all of the new Congress's attitudes towards science and maths education are entirely beneficial. For example, Congress might have gone too far in focusing almost exclusively on equity issues in education. There's no question that the United States needs to target some of its education efforts towards schools in poorer school districts and under-represented minorities, especially as those groups become a larger proportion of the population. Federal programmes have an obligation to compensate for funding disparities in the nation's locally based education system, and

indeed that's what most of the Department of Education's programmes are designed to do.

But the new law targets almost all of its numerous programmes towards schools in poor areas. That seems to go too far. Better teaching and other ways to interest more students in science are needed in a wide range of schools. The federal government needs a portfolio of programmes that engage with the full range of school districts — an approach that could be hampered by the new statute.

The new law does take a portfolio approach in one sense — it creates a dizzying array of sometimes overlapping programmes rather than setting priorities. That's because the only way to get the law passed so quickly — seven months after the start of the session — was to include almost every programme sought by interested legislators.

For example, the law authorizes at least four new pre-college programmes involving the national laboratories — pushed by Republican Senator Pete Domenici and Democratic Senator Jeff Bingaman, who, unsurprisingly given their home turf of New Mexico, are advocates for the labs. Encouraging the labs to interact more with schools makes good sense. But is it a good use of DoE funds to help states establish or expand schools that specialize in maths and science education that would then work with the labs? Probably not.

Interestingly, the DoE spending bills do not seem to set aside any money for education programmes, although they could still be funded out of general appropriations for DoE science, which are slated to rise next year.

As a whole, though, the new act is a plus for science education, especially given how uneasy Congress still feels about federal involvement in education 50 years after Sputnik. Back then, the National Defense Education Act opened with a statement that it was not intended to signal a federal takeover of education. And the new law explicitly states that the Department of Education "shall not endorse, approve or sanction any mathematics curriculum designed for use in any school" — even as it authorizes the department to fund projects designed to improve maths instruction in the neediest schools by using new instructional materials. ■

David Goldston is a visiting lecturer at Princeton University's Woodrow Wilson School of Public and International Affairs.

BUSINESS

Copycat consolidation

As more blockbuster drugs come off patent, generic drugmakers face a changing landscape.

Meredith Wadman looks at their strategies for survival.

One year ago, Mylan Laboratories was the world's ninth-largest generic drugmaker, a thoroughly American company with American clientele and American operations.

Today, on the heels of two major acquisitions, the Pennsylvania-based company has vaulted into the ranks of the top three global generics players. In January, Mylan bought a controlling interest in Matrix Laboratories, an Indian company that makes a huge range of the active ingredients in generic drugs. Then in May it agreed to spend a cool US\$6.7 billion to buy the generics division of Germany's Merck KGaA, which sells copycat pills in more than 90 countries. Once the deal closes on 1 October, Mylan will become the world's third-largest generics maker by sales, after Teva Pharmaceutical Industries in Israel, and Sandoz, the generic unit of Swiss brand-name drugmaker Novartis.

In the process, Mylan will have built for itself "one of the most global platforms there is," says Tommy Erdei, executive director of the global health-care group at UBS Investment Bank in London. "This is the kind of structure that the generic pharmaceutical industry is going towards."

Indeed, seven major mergers in the past two years alone have marked a wave of consolidation in the generics industry — companies that make and sell cheaper copies of off-patent pills. That wave is reshaping the industry as a highly competitive battleground with scores of Davids, a handful of emerging Goliaths and numerous mid-sized companies caught in between.

Going global

Last year, Teva's acquisition of Ivax in Florida created an uncontested global leader that now controls 20% of the US generics business. Last October, Barr Pharmaceuticals in New York shelled out \$2.5 billion for the Croatian generics maker Pliva in Zagreb, giving it instant access to sales outlets in 30 countries — and to Pliva's low-cost manufacturing facilities in Croatia and Poland. Meanwhile, Actavis,



Mylan Laboratories is about to become the third-largest generic-drug company.

which a decade ago was an obscure Icelandic company, has acquired more than 25 companies in seven years, in the process extending, amoeba-like, into 40 countries — and becoming the sixth-place global player last year.

Even before the acquisitions by Barr and Mylan, the wave of consolidation had had a palpable effect in the United States. Last December, the top four generics companies controlled 56% of the market, compared with 35% ten years ago, according to the industry information firm IMS Health in Connecticut. And it's a bigger pie they're controlling: a groundswell of patent expirations has pushed up the generic share of prescriptions in the United States from 47% to 63% since 2000.

"You have to launch a new product into all the key markets."

Driving the consolidation is the major players' desire to have a global presence in an industry in which the competition has become brutal. In the United States, it's not uncommon for 15–20 companies to start marketing generic versions of a drug once patent protection has expired. Many of them are small players, scrabbling for tiny slices of a market dominated by Teva and other giants. But because they're small, they're often willing to slash prices to recoup their investment, driving prices down much further and faster than the major players would do if left to their own devices.

"To compensate for all the price erosions and to keep growing, it's not sufficient to launch [a new product] in two or three countries," Robert Wessman, the chief executive of Actavis,

told a conference of the Generic Pharmaceuticals Association in Washington DC this month. "It's obvious that you have to launch it into all the key markets in the world." And since some of those markets — in countries such as Italy, Spain and France — aren't yet as well penetrated as the United States, the impetus to get bigger and more global makes mergers an obvious strategy.

Still, some, including Wessman, say that the merger mania is driving companies to overpay for acquisitions: he fingers both the Mylan–Merck and Barr–Pliva price tags as too high — but then, he was outbid by Barr in the chase

for Pliva. Bruce Downey, Barr's chief executive, says he has no regrets. "We got great value because of all the assets we acquired in the acquisition," he says. "I would do it again."

A fast-moving industry

Not to be left behind, some Indian companies are turning the tables and acquiring US makers of generic drugs. "Indian players are really building up a presence in the United States," says Doug Long, vice-president of industry relations at IMS Health. For instance, Mumbai-based Wockhardt, already the largest Indian player in Europe, was last week reported to be bidding for one of the companies bidding for New Jersey-based Par Pharmaceutical. Others are going global: consider, for example, the Indian firm Ranbaxy Laboratories, which acquired generics companies in Italy, Romania and Belgium all in one week early last year.

So will the natural selection of the business world leave only a few generic giants standing in 10 or 15 years time? "That's what a lot of people believe," says Long.

Erdei reckons that small players in highly targeted niche markets such as oncology and dermatology will survive. "The middle is the group we're most worried about," he says, because they're too big to make a niche strategy work for them but too small to compete with the likes of Teva and Mylan. These middle companies need to think about how to transform themselves into Goliaths, he says: "Because if they don't do it proactively, someone else will do it for them."



BRINGING THE PAST TO LIFE

Can a stage spectacular based on a TV documentary bring science to life and please the punters too?
Brendan Maher joins a palaeontologist to watch the dinosaurs walk.

As the seating in the arena slowly fills up the children can't hide their giddy anticipation. Neither can Ken Lacovara, chattering away about dinosaurs and digs. Admittedly, he has a soul patch on his chin, a beer and a couple of graduate students, which marks him out from the majority of enthusiasts here to see *Walking with Dinosaurs: The Live Experience*. But there's no mistaking the kinship between the professor of geology and palaeontology from Drexel University, Philadelphia, and the children around him. "Everybody I know in the field wanted to do this since they were very young," he says, looking around. "You never know what future scientists might be in the audience."

The stage show now touring America — watched by *Nature* and Lacovara at Philadelphia's Wachovia Spectrum sports arena in August — was inspired by the BBC documentary series, which, according to Lacovara, set a gold standard for edutainment (see 'Origin of a Species'). Using computer animation, animatronics and the authoritative tones of the actor Kenneth Branagh, the series told the tale of the dinosaurs' 160-million-year

lease on Earth. The programme was stunning to watch yet stuck close enough to scientific understanding not to upset an expert. At least, not Lacovara. The \$20-million stage show has a lot to live up to, and Lacovara's excitement — especially over the prospect of a fully fleshed-out brachiosaurus — is tinged with scepticism. How well can a theatrical presentation relay natural history? Will spectacle triumph at the expense of information?

The house lights dim and a booming voice reminds the crowd that, as cell phones and pagers didn't exist 65 million years ago, they should be turned off. Score one for realism. 'Huxley', a palaeontological P. T. Barnum strides on to the stage to serve as our guide and scale bar. He bends down by a nest from which plateosaurus hatchlings emerge, in the form of squirming green hand-puppets. The eggs, Lacovara notes, aren't shaped quite right — too "chickeny" — but his criticism is cut short by lilliensternus. A two-metre-tall carnivore enters the arena. Actually, it's a suit worn by an

actor with the mettle to carry 40 kilograms of foam, lycra and animatronics on his back and yet still look nimble. Once you learn to ignore the craftily camouflaged extra set of human legs, it's pretty convincing.

In the interest of conflict, a full-grown plateosaurus appears next, eager to defend its babies. This large dinosaur is a puppet, operated by three people. One drives a slim car camouflaged beneath the dinosaur; two others are in a control room moving its neck, tail, jaws and the like by manipulating a smaller version, evocatively known as a voodoo rig. Prosauropod and predator settle into a carefully choreographed stand-off. The slow, deliberate tempo

is the result of much trial and error, Matthew McCoy, the head of puppetry, later explains. McCoy tells the tale of a tragic show in Sydney, Australia, in which *Tyrannosaurus rex*'s head fell off after a tight turn at high speed. The audience was sympathetic, he says with some gratitude, but the troupe learned its lesson.

In addition to slowing down the action in

"Prosauropod and predator settle into a carefully choreographed stand-off."

T.S. WARREN/AP

Origin of a species

Would you expect to see your first dinosaur on a dig, in a museum — or at a cocktail party? In 2003, William May, the artistic director of the stage show *Walking with Dinosaurs: The Live Experience*, invited a collection of socialites and friends to his workshop, Number 12, Elizabeth Street, Kensington, Melbourne, Australia, a converted wool warehouse and tannery decorated with trolls, fairies and other larger-than-life puppets from May's theatrical exploits. As they were sipping wine and nibbling canapés, a 14-metre *Tyrannosaurus rex* burst in. Fashioned from a mechanical crane with garden-hose ribs and Lycra skin, the first prototype for the show now touring America debuted in style.

When the BBC first approached May and his business partner Malcolm Cooke earlier that year with the challenge of creating a stage version of its acclaimed TV documentary, May knew that no cartoonish facsimiles would do. The only way forward was to use full-scale models, a technical challenge that launched him into two years of research. From the size, balance and gait to the skin, teeth and eyes, everything had to be not only as technically accurate as possible, but absolutely stunning to the audience. "Can you bring dinosaurs back to life in the most realistic form that could tour the world, beat the most cynical child, beat the scientist, beat the reviewers? It's like going to the Moon."

Brooklyn-born May says he's always lived by the motto that he could make anything happen if he put his mind to it. Practising ballet in secret in his bedroom as a child, he grew to become one of the youngest producers of musicals on Broadway. There, he learned to bring imagination to life through puppetry and animatronics.

Dealing with scientists was a change of pace, he says. "I love science. You get to meet the most brilliant minds." But as the dinosaur's drinks-party debut demonstrated, he's a showman at heart.

"The future of theatre won't come from song and dance but from pushing boundaries. Audiences are very discerning now; they play computer games and watch movies with computer-generated imagery. This isn't a film, or a trick, this is an emotional, real experience."

Julie Clayton and Ruth Francis

Listen to William May and Ken Lacovara discuss dinosaurs on this week's Nature podcast (www.nature.com/podcast)



Head to head: palaeontologist Ken Lacovara in a montage with dinosaurs used in the performance.

later shows, the team built a spare *T. rex*. With 15 dinosaurs taking the stage every night, losing one wouldn't necessarily stop the show, but turning up in front of hundreds of children without a working *T. rex* just doesn't cut it. "We might as well just go home," says McCoy. There are contingency plans for other disasters, too. Had lilliensternus been toppled by plateosaurus's heavy whiplike tail, for example, he would have needed help getting back up. That, says McCoy, is when they send in the dinosaur clowns.

No disasters strike in Philadelphia. After several minutes of a mock battle and trotting about, lilliensternus and plateosaurus dutifully leave the stage, and Huxley eases the crowd through geological time into the Jurassic period. Bright inflatable plants explode around the stage. Lacovara gleefully elbows one of his students; he thinks giving a round of applause to the Jurassic just for starting is pretty amusing.

With the Jurassic period come the brachiosaurs. The young one is perhaps two storeys tall; the adult, more than ten metres. They make an impressive pair as they stretch their long necks deep into the stands, delighting the audience. As one of them almost lays its head in Lacovara's lap, he notes that a full-grown adult would have been a bit taller, but he's still impressed. The palaeontologist on stage rattles off statistics about the beast, which may have weighed as much as 40 tonnes. The expert in the stands notes, with a mischievous grin, that the titanosaur his group is excavating in Patagonia weighed 60.

But despite a little professional one-upman-

ship, Lacovara likes the show. During the intermission, he confers with his students, who agree that it doesn't pull any educational punches. "It's just packed full of content," Lacovara says, noting that it introduces concepts such as deep time, plate tectonics, climate change and evolution: all ingredients, he says, that presented too dryly would spell certain death.

This concept brings us ineluctably to the show's finale. In the climactic Late Cretaceous, rife with volcanic drama (cue the light show), a *T. rex* mother and son take the stage triumphantly. After some play-fighting with ankylosaurus and torosaurus, they turn their attentions to the crowd. While baby *rex*, another actor in a heavy dino suit, mugs for the crowd, mum is scaring the life out of them. A blonde boy just behind Lacovara chats nervously with his father about the seating arrangement as the *T. rex* approaches. "No Daddy, don't tell him I'm here."

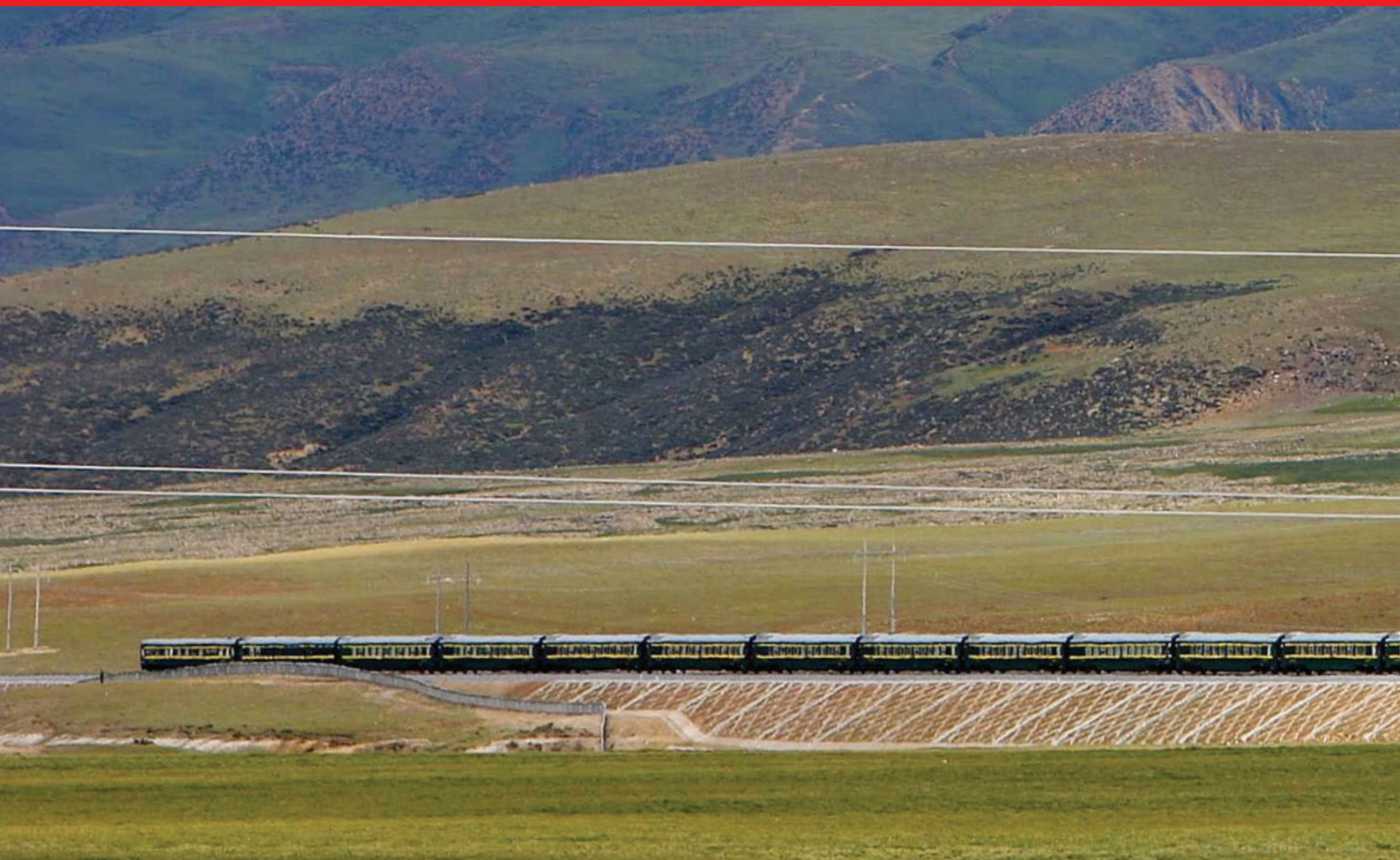
A bright strobe with booming audio represents the extraterrestrial *coup de grâce* at the end of the Cretaceous. The dinos exit, the plants deflate, and bows are taken. The crowd drains from the Spectrum, and the children are laden with bright and blinking palaeoparaphernalia. After three more shows, the crew will pack its 27 truckloads of equipment and move on to the next stop. Lacovara is beaming, satisfied with the production's portrayal of the work he does. Dinosaurs, he says, are "a gateway drug for the sciences". A lot of kids scored tonight. ■

Brendan Maher is a features editor at Nature.

"Without a working *Tyrannosaurus rex* we might as well just go home."

— Matthew McCoy

B. MAHER



Riding on the roof of the world

China's railway to Tibet is an engineering marvel or an environmental menace — or perhaps both. **Jane Qiu** takes a ride to find out.

It is a glorious, crispy cold summer morning on the Qinghai–Tibet plateau, and already my lips are turning purple. At more than 4,500 metres above sea level, the air is thin and I can feel a light headache coming on. In front of me a pair of rail tracks stretch into the distance, looking as thin as silver threads as they negotiate a landscape filled by expansive glaciers and mountains with needle-sharp peaks.

The rattling sound of an approaching train jolts me to alertness and in no time it roars past us at 100 kilometres an hour, passengers waving gleefully from its windows. I'm painfully aware that the pressurized air behind those windows offers a great deal more oxygen than the stuff I'm breathing in the open air. "You will be fine," my Tibetan host, Tsega, pronounces after inspecting me up and down a few times. I feel obliged to trust his judgement.

The railway in front of me is the world's highest, and is a 1,142-kilometre stretch that connects Golmud, in China's Qinghai province, with Lhasa, capital of the Tibet Autonomous Region. It is also widely regarded as one of the great engineering achievements of the world. Its course crosses four mountain chains and five

major rivers — with nearly all of it more than 4,000 metres above sea level. Successive Chinese governments, keen to tighten their grip on Tibet, have dreamed of such a railway for nearly half a century, and more than 20,000 workers laboured for 5 years to complete the project at a cost of 33 billion yuan (about US\$4 billion).

Now, more than a year after it opened, the railway remains a source of bitter controversy. Supporters say the scheme will bring major opportunities to China's underdeveloped west. Critics fear that China will use it to assert control over a contested border region, and to exploit its natural resources. The railway's long-term impacts on the plateau — direct and indirect — remain unknown, and I have embarked on a journey between Golmud and Lhasa — sometimes on the train, sometimes off it — to seek answers. The trip reveals how well the engineering is standing up after a year's worth of exposure to the harsh seasons; how the attempts to minimize its environmental

effects have fared; and what the railway means to ordinary Tibetans.

Earthquakes and permafrost

Tsega, who runs the Kekexili nature reserve from the office in Golmud, has kindly agreed to take me along on one of his regular surveillance trips to the reserve. After driving for 100 kilometres, we approach the gigantic Kunlun fault, a 430-kilometre surface rupture, ripped open by a magnitude-8.1 earthquake¹ on 14 November 2001. The scar from that quake is clearly visible in the landscape, and sends

a shiver down my spine as we drive across it.

The earthquake didn't cause much damage to the railway because construction had just started, says Wang Lanming, director of the Lanzhou Institute of Seismology, China Earthquake Administration, in Gansu province. "But it was certainly a wake-up call for those of us involved in advising the engineers of the region's seismic activity," he adds. Two

"It would be very dangerous if passengers were exposed to the harsh conditions of the Qinghai-Tibet plateau."
— Wang Lanming



seasonal freeze–thaw cycles. Construction is even harder with the mix of permafrost types on the Qinghai–Tibet plateau. The railway traverses 275 kilometres of warm permafrost, 221 kilometres of ice-rich permafrost, and 134 kilometres of a mix of both. “This is the worst combination for any permafrost engineering project,” says Max Brewer, a permafrost researcher at the University of Alaska in Fairbanks.

SIPA/REX FEATURES

Keeping cool

A group of researchers, led by Cheng Guodong of the Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, in Lanzhou, has developed a series of measures to cool down the railbed. All the approaches capitalize on using the benefits of the cold environment, and try to keep the railbed as naturally cold as possible².

In most places, for instance, the railway embankment is raised by between 2 metres and 10 metres, helping insulate the ground from the heat created by the tracks. The slopes of the embankment are also covered with a layer of crushed rocks — a technique inspired by reports that permafrost can occur beneath blocky and coarse materials even when the air temperature is well above 0 °C. “This is the first time a large-scale project has used the technique as one of its primary solutions,” says Zhang Tingjun, a permafrost researcher at the National Snow and Ice Data Center in Boulder, Colorado. It seems to be working; tests suggest that the crushed layer keeps temperatures up to 2 °C cooler, and in some places, the permafrost has even increased in volume and pushed up into the embankment structure.

As we drive along, passing Tibetan wild asses

other quakes of magnitude 7.5 or greater have shaken the plateau in the past century, and seismic activity poses one of the major threats to the new railway.

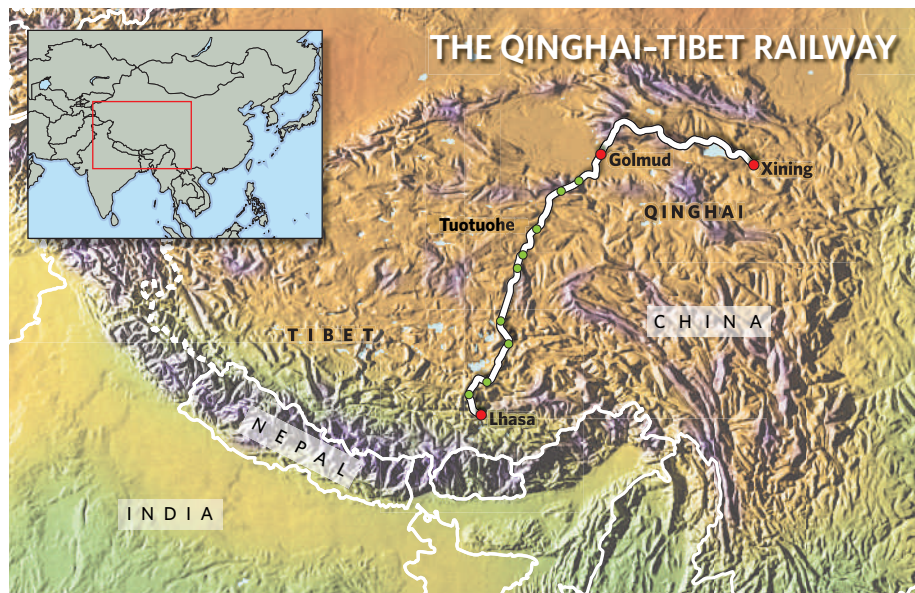
The question is what to do about it. Railway engineers have avoided building stations, tunnels and bridges in areas containing active faults, and these main structures are equipped with anti-quake safety measures such as additional steel reinforcement bars to keep the concrete from cracking during an earthquake. But the tracks themselves do not have any extra reinforcement. “The cost would be astronomical to install anti-quake measures all through the route,” says Wang.

However, researchers and engineers are in talks with the ministry of railways about installing an earthquake-detection system along the tracks on the plateau. The project, estimated to cost tens of millions of yuan, could, in theory, notify the train of a substantial earthquake in progress by detecting the primary seismic waves propagating from a quake, then bringing the train to a halt before the damaging secondary waves arrive. As Wang notes, stopping a train before it derailed in a quake is particularly important for high-altitude railways: “It would be very dangerous if passengers were exposed to the harsh conditions of the Qinghai–Tibet plateau,” he says.

The harsh weather here also calls for innovative approaches to railway engineering. As we crest the Kunlun Range, ahead of us

we see how the rail tracks disappear into a 1.6-kilometre-long tunnel. “We are entering a region rich with warm permafrost,” Tsega explains. The temperature of warm permafrost stays within a couple of degrees below freezing, and its top layer thaws more readily in the summer than that of normal permafrost, making the ground more unstable.

Regular permafrost is bad enough; engineers have to raise the railbed or use insulating materials to keep the tracks from warping during



grazing by the road, I spot other gadgets along the tracks for cooling the embankment. The most prominent ones are 'thermosyphons' — a series of thin metal tubes standing upright on the railbed shoulders every 3 metres, which use evaporative cooling to transfer heat from one end to another. Thermosyphons are costly to install and maintain, but the Chinese government is not shy in lavishing on such gimmicks, with 18,200 of them installed in 'high-risk' permafrost areas.

Elsewhere, engineers buried ventilation pipes through the embankments, which can also cool the temperature by up to 2 °C. And some of the south-facing slopes hold giant shading-boards. Over the first winter, these allowed shaded areas to freeze solid while non-shaded embankments remained unfrozen.

More than 200 sensors constantly monitor ground temperatures on both sides of the railbed over permafrost regions, in a system costing 40 million yuan. One year and a seasonal cycle on, researchers suggest that the permafrost regions are recovering from the disturbance caused during the construction of the railway. But there have been several cracks along the track in some regions, which may be due to distorted foundations. Cracks are repaired as soon as they are discovered by the crew regularly monitoring the track. Zhang Luxin, chief engineer for construction and maintenance of the railway, would not disclose the total number of cracks since it began operating. But he says that all of the cracks so far have been superficial and have had no impact on the railbed.

Brewer says he is not alarmed by such incidents. "It would be naive to expect that construction of this scale could be free of problems," he says. Indeed, it has been more than a century since Siberia's first railway was built, but frost-damage still affects 30% of the tracks passing across permafrost regions, according to a recent survey conducted by Russian engineers.



Cooling pipes: ventilation ducts through the embankment keep the ground temperature up to 2 °C lower.

Some experts, however, fear that the problems could get worse in a warmer climate. Research shows that global warming has affected the ground temperature on the plateau to a depth of 40 metres; permafrost is degrading as temperatures rise and the 'active' layer that freezes and thaws every year gets thicker³. According to a recent report by the Chinese Academy of Sciences, 10% of the permafrost regions on the plateau have degraded in the past 20 years.

Before 2001, when the railway construction started, the consensus was that the air temperature on the plateau would increase by 1 °C over the next 50 years. Scientists now think that it may rise by 2.2–2.6 °C during that time. Cheng, however, says that engineers could adapt the railways to additional temperature increases by adding more of the cooling measures to the railbed's slopes.

Lack of oxygen sends excruciating pulses burrowing in through the top of my head and deep into my brain. I turn to the view outside, seeking comfort in the widening, sparsely vegetated valley, sprinkled with streams, rivers and enormous alluvial fans. In the distance, my gaze comes to rest on a gigantic overpass perching high above the plain. "That's

the famous Qingshuihe Bridge," Tsega says, noticing my sudden enthusiasm. The 11.7-kilometre-long overpass, consisting of 1,367 piers and costing 24 million yuan to build, straddles the most unstable stretch of permafrost along the railway. The overpass removes direct contact between permafrost and the tracks — the engineers' last resort for protecting fragile ground.

Effect on migration

Tsega explains that the bridge is also one of 33 passageways designed to enable wildlife on the plateau to go under the

arches to get from one side of the railway to the other. Chiru, endangered Tibetan antelopes known for their speed and stamina, are shy creatures and particularly prone to disturbances. Every June, more than 3,000 pregnant chiru travel hundreds of kilometres westward from the Three-River Headwater nature reserve to give birth in the Kekexili reserve, then migrate back with their young in August. Both the Qinghai–Tibet railway, and the highway that follows the same route, cut right across their migratory paths.

Experts are divided on the subject of how effective such wildlife passageways are. According to a government report released ahead of the first anniversary of the line's operation in July, most of the passageways are used by plateau animals and the railway has had no impact on the Tibetan antelopes in those areas. Other research is less positive. Contrary to reports by the state-run Chinese media, as many as 1,500 antelopes couldn't make the crossing in 2003 and had to give birth locally, says Yang Xin, president of Green River, a non-

"It would be naive to expect that construction of this scale could be free of problems."
— Max Brewer



The chiru is an endangered antelope found only on the Tibetan plateau.



governmental organization based in Chengdu, Sichuan province. This happened despite the fact that railway workers suspended construction and cleared out of the sites for a couple of days during the animals' peak migration period.

Although antelopes are able to cross the railway, research by Yang and Su Jianping, a zoologist at the Northwest Institute of Plateau Biology, Chinese Academy of Sciences, in Xining, Qinghai province, suggests that antelopes don't use most of the passageways. Before the railway construction began, antelopes crossed the highway between the Kunlun Pass and Wudaoliang at multiple locations along a 100-kilometre stretch. Now, a large majority of them funnel under the 200-metre-long overpass at the Wubei Bridge, just north of Wudaoliang.

It's not yet clear what these changes in the animals' migration habits might mean in the long term. The researchers conjecture that some of the new overpasses might be too low or narrow for animals to feel comfortable, which pushes them into new migration routes. Yang, for instance, has seen herds of antelope approaching the Chumaer River — which used to be a key crossing site — then hesitate, turn away and go south to make the crossing near Wudaoliang.

Knowing why some overpasses seem to work better than others could be valuable for future construction work, says Su — but only long-term monitoring can find that out. "It would be surprising if the highway and railway didn't affect movements of plateau animals," he says. "It's unclear how this could affect their other behaviours or their survival as a species." Yang Qisen of the Beijing-based Institute of Zoology — leader of the team that studies the railway's

Tang Gu La, at 5,068 metres above sea level, is the world's highest train station.

impact on plateau wildlife — declined *Nature's* request for an interview.

Throughout the journey, the plateau is otherwise little marred by construction debris or the landscape scars usually associated with major projects. Railway planners developed a 'green policy' to ensure protection for soil, vegetation, animals and water resources along the route⁴. Nearly 4% of the 33 billion yuan spent on construction was allocated to restoration of ecosystems and environmental protection, according to Hao Qingjun, an official at the Xining-based Qinghai Environmental Protection Agency. Where possible, the railway was rerouted around sensitive natural zones such as wetlands. In other places, much of the turf that was dug up during the construction phase was preserved and then replanted once that section of the railway was completed.

Accumulating rubbish

The sun has almost set by the time we pull off the road at the reserve's protection station near Tuotuohe for the last stop of the day. The lingering light sets the snow-capped mountain peaks ablaze, while long blue shadows creep down from their dark stony slopes.

Tuotuohe, a main town along the railway, lies in the region that is the headwater of three of Asia's largest rivers — the Yangtze, Yellow and

Lancang (Mekong) rivers. Standing at the Tuotuohe Bridge, the first bridge over the Yangtze, I recall how Yang told me that the town is also the first rubbish dump along the river. The highway and railway are bringing more people and industrial products to the plateau. They have also led to more roads and other constructions, such as hotels and restaurants, along the route. "Those constructions do not have the same kind of resources as the railway project," says Yang, referring to the lack of money and manpower. "So environmental protection is not on the agenda."

Few towns along the route have the capacity to treat their accumulating rubbish, so it is left to moulder in open containers on the street and along the river banks. This is not only a source of disease, says Yang, but it also pollutes the environment and endangers wildlife. In addition, the introduction of commerce into the plateau life has led to an increasing need for more livestock, such as yaks, cows and sheep, which nomads can exchange for industrial products. "Over-grazing is a serious problem on the plateau," he says. In the Three-River Headwater region, some nomads have to go as high as 5,500 metres to find fresh pasture for their animals.

Change in Tibet

It soon turns pitch-dark and the temperature drops sharply. In the protection station we sit around a stove, enjoying the waves of warmth it radiates. A television is showing news on Phoenix TV, a popular Hong Kong channel. There is a computer in the corner with an internet connection. Tashi, a young Tibetan man who looks after the station, brings us supper: bread, yoghurt and Tibetan tea, along

with some dried beef Tsega brought with us. Over the meal, I ask Tashi how the railway has affected his life and that of his family who are nomads in the region. "Our life is still the same," he says after a moment of hesitation, a sense of indif-

ference palpable. None of his family or any of their friends sees a need for taking the train. "We are nomads."

His indifference, and that of many others living in rural Tibet, is in stark contrast to China's claimed purpose of the railway: to promote economic development. On the surface, Tibet is doing well, and the railway has certainly helped; during the first 10 months of its operation, trade between Tibet and outside regions increased by 75%, to 2.6 billion yuan. Tibet will receive an estimated 3 million tourists this year and, by 2010, the number is expected to reach 5 million.

"Our life is still the same. We are nomads."

— Tashi



Pleased with the railway? The reaction in Tibet to the railway project has been mixed.

The billions of yuan that Beijing has poured into Tibet flow almost exclusively along the handful of roads in the region, pooling in the towns along the way and finally ending up in three major cities — Lhasa, Shigatse and Chamdo, says Kabir Heimsath, a Lhasa-based anthropologist at St Cross College, University of Oxford. The towns and cities are the few places in which serious development has taken hold and, consequently, are also the only places where Han migrants — mostly itinerant workers marginalized by the same system in other parts of China — and businessmen form a majority. This increasing disparity between the urban rich and the rural poor — a general problem of China's economic development — means that most Tibetans are left behind, unable to reap the benefits of modernization, he says.

Even in towns and cities, Tibetans are in danger of being marginalized. They find it increasingly difficult to compete in the job market against skilled and more business-oriented Han, who, privileged by ethnicity and language, make up more than 92% of the population of China. Intentionally or not, Tibet has seen a surge of Han migrants, which is further boosted by the railway. And emerging with this wave of migration is the cultural imprint of Han Chinese on many

aspects of Tibetan life — even religion, as an increasing number of Tibetan monasteries are funded and built by Han Chinese.

On the train

A few days after visiting Tuotuohe, I board the train from Golmud to Lhasa for the uninterrupted experience of the entire train journey. The coaches are completely sealed so no rubbish can be thrown out and have a wastewater-storage system. They have windows with ultraviolet filters to keep out the sun's glare, and their underbellies are enclosed to protect wiring from snowstorms and sandstorms. The oxygen level in every coach is well-adjusted, and additionally, passengers can plug in to the oxygen supply anywhere on the train. Every coach is equipped with digital displays that show continuous updates of the train's elevation and speed, the outside temperature and the distance to the next station.

Travelling at about 100 kilometres an hour, the trip is surprisingly smooth. I have the opportunity to review the scenery — the magnificent Mount Yuzhu reaching up to 6,178 metres, the turquoise Namucuo Lake shimmering under the unearthly Tibetan light, nomads dressed in colourful robes, turrets with red, blue and green prayer flags fluttering near a Communist flag. Fourteen hours later, at the

end of an uneventful journey, the train pulls into the great vault of Lhasa's new railway station.

The next morning I wander through the city's various districts. Apart from the Tibetan quarter centred on the holiest Jokang temple, most of the shops, hotels and restaurants are run by Han Chinese. I come across Bhuchu, a Tibetan pilgrim from a village in Sichuan province. He has prostrated himself on the ground after every step for more than 1,000 kilometres over the past 6 months. Now only steps away from Jokang, he is radiant with joy, hope and a deep sense of heightened existence.

I finally arrive at the Potala Palace, the historic landmark of Lhasa, which used to be home to successive Dalai Lamas. The sight of the gigantic red-and-white structure perching serenely on Potala Hill is simply breathtaking. Marvelling at its architectural splendour and spiritual richness, I can't help but wonder what the railway will bring to Tibet and its people.

Already, China is planning to extend the railway even farther, to Zhangmu to the west and to Dali to the east. Another extension would link Shigatse with Yadong, near the China-India border. What will this expanded network mean for the fragile plateau? Here in the rarefied environment of Lhasa, the question hangs shimmering in the air. ■

Jane Qiu is a freelance writer based in Beijing and London.



Untreated rubbish has accumulated in places along the Yangtze river.

GREEN RIVER

SIPA/REX FEATURES

1. Xu, X., Chen, W., Ma, W., Yu, G. & Chen, G. *Seismol. Res. Lett.* **73**, 884–892 (2002).
2. Cheng, G., Sun, Z. & Niu, F. *Cold Reg. Sci. Technol.*, doi:10.1016/j.coldregions.2007.02.006 (2007).
3. Cheng, G. & Wu, T. *J. Geophys. Res.* **112**, F02503 (2007).
4. Peng, C. *et al. Science* **316**, 546–547 (2007).

Ocean pipes could help the Earth to cure itself

SIR — We propose a way to stimulate the Earth's capacity to cure itself, as an emergency treatment for the pathology of global warming.

Measurements of the climate system show that the Earth is fast becoming a hotter planet than anything yet experienced by humans. Processes that would normally regulate climate are being driven to amplify warming. Such feedbacks, as well as the inertia of the Earth system — and that of our response — make it doubtful that any of the well-intentioned technical or social schemes for carbon dieting will restore the status quo. What is needed is a fundamental cure.

The oceans, which cover more than 70% of the Earth's surface, are a promising place to seek a regulating influence. One approach would be to use free-floating or tethered vertical pipes to increase the mixing of nutrient-rich waters below the thermocline with the relatively barren waters at the ocean surface. (We acknowledge advice from Armand Neukermans on engineering aspects of the pipes.) Water pumped up pipes — say, 100 to 200 metres long, 10 metres in diameter and with a one-way flap valve at the lower end for pumping by wave movement — would fertilize algae in the surface waters and encourage them to bloom. This would pump down carbon dioxide and produce dimethyl sulphide, the precursor of nuclei that form sunlight-reflecting clouds.

Such an approach may fail, perhaps on engineering or economic grounds. And the impact on ocean acidification will need to be taken into account.

But the stakes are so high that we put forward the general concept of using the Earth system's own energy for amelioration. The removal of 500 gigatonnes of carbon dioxide from the air by human endeavour is beyond our current technological capability. If we can't 'heal the planet' directly, we may be able to help the planet heal itself.

James E. Lovelock*, **Chris G. Rapley†**

*Green College, University of Oxford, Woodstock Road, Oxford OX2 6HG, UK

†Science Museum, Exhibition Road, South Kensington, London SW7 2DD, UK

H-index: however ranked, citations need context

SIR — The h-index (the number n of a researcher's papers that have received at least n citations) may paint a more objective picture of productivity than some metrics, as your News story 'Achievement index climbs the ranks' (*Nature* **448**, 737; 2007) points out.

But for all such metrics, context is critical.

Many citations are used simply to flesh out a paper's introduction, having no real significance to the work. Citations are also sometimes made in a negative context, or to fraudulent or retracted publications. Other confounding factors include the practice of 'gratuitous authorship' and the so-called 'Matthew effect', whereby well-established researchers and projects are cited disproportionately more often than those that are less widely known. Finally, bibliometrics do not compensate for the well-known citation bias that favours review articles.

Michael C. Wendl

Washington University Medical School, 4444 Forest Park Boulevard, Box 8501, St Louis, Missouri 63108, USA

H-index: age and sex make it unreliable

SIR — The h-index seems to be breaking away from the bibliometric pack, in the race to become a favoured measure of scientific performance ('Achievement index climbs the ranks' *Nature* **448**, 737; 2007). However, if the h-index is to become an assessment tool commonly used by university administrators and government bureaucrats, those using it should be aware of its pitfalls.

As noted in your News story, tallying how many papers a researcher publishes (their productivity) gives undue merit to those who publish many inconsequential papers. But at least for ecologists and evolutionary biologists, the h-index is highly correlated with productivity ($r = 0.77$; see C. D. Kelly and M. D. Jennions *Trends Ecol. Evol.* **21**, 167–170; 2006).

This is worrisome, because the h-index is easily misconstrued as an equitable measure of research quality. We offer two examples.

First, female ecologists and evolutionary biologists publish fewer papers than their male counterparts, and they have significantly lower h-indices. Should administrators therefore conclude that men are better researchers? No. The gender difference vanishes if we control for productivity. It seems unlikely that this phenomenon is restricted to ecology and evolution.

Second, the h-index increases with age and using the ratio of the two can be problematic. Therefore, reliably comparing the performance of younger researchers with older ones is difficult.

Clint D. Kelly, **Michael D. Jennions**

School of Botany and Zoology, Australian National University, Canberra, ACT 0200, Australia

Conservationists could slip up in oil-palm enterprise

SIR — In a novel 'farming with the enemy' approach, Lian Pin Koh and David S. Wilcove (*Nature* **448**, 993–994; 2007) suggest that non-governmental organizations (NGOs) should cash in on the lucrative oil-palm industry and buy small tracts of existing plantations to generate revenue for nature reserves. Such a scheme has many potential complications and ramifications.

According to Koh and Wilcove's figures, an NGO would require \$62.5 million to purchase 5,000 hectares of mature oil palm; the same sum of money would buy 11,000 hectares of forested land outright. It would take six years to recoup this investment through revenues from the plantation. Only then could forested land be bought, at 1,800 hectares a year, reaching a total of 11,000 hectares 12 years after the oil-palm purchase. Given the high rates of deforestation in southeast Asia, it seems more prudent to simply acquire forests now.

NGOs may spread themselves too thin if they engage in a large business that will divert considerable manpower, time and expertise away from threatened-species recovery programmes and public outreach. In addition, acquisition of oil-palm plantations by NGOs may exacerbate current rates of land conversion, by reducing the number of hectares available for other private plantation owners. And we fail to see how the proposed scheme will prompt companies to commit to good management practices, such as those advocated by the Roundtable on Sustainable Palm Oil (www.rspo.org). Instead, oil-palm monoculture could be legitimized as 'green' by an NGO's association with companies that are required to do nothing more than assist in starting up their plantations.

We propose instead that NGOs actively engage oil-palm companies to fund reserves, in return for environmentally friendly accreditation and help in implementing guidelines for sustainability.

Although conservationists must find creative ways to raise funds sustainably, doing so through an industry that has negative impacts on biodiversity may ultimately be counterproductive.

Reuben Clements, **Mary Rose C. Posa**

Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Republic of Singapore

Contributions to Correspondence may be submitted to correspondence@nature.com. They should be no longer than 300 words. Published contributions are edited. We welcome comments at [Nautilus](http://blogs.nature.com/nautilus) (<http://blogs.nature.com/nautilus>).

BOOKS & ARTS

Research funding in the twenty-first century

Are universities and market forces compatible?

Science for Sale: The Perils, Rewards and Delusions of Campus Capitalism

by Daniel S. Greenberg

University of Chicago Press: 2007.

288 pp. \$25

Michael M. Crow

Daniel Greenberg is widely considered the premier journalist of science policy, having written extensively on the subject over the course of its 60-year evolution in the United States. *Science for Sale* is his latest offering. It provides an intriguing, if idealistic, review of the issues surrounding the funding of science in the twenty-first century. Greenberg posits that science was once, and should be again, driven by the pure curiosity of scientists and not by motives influenced by the stress of external funding and the negative forces of capitalism. Unfortunately, science past did not really exist in the way he spends so much time describing in the book.

Greenberg's idyllic views — in particular that the academic scientist and the university are best motivated by curiosity alone — are interesting. But they run counter to history, to how organizations operate and, perhaps most importantly, to the understanding that 'the university' itself is an idea, not an ideal or an ideology.

In fine journalistic style, the author makes his case that the modern university, with its quest for resources, has drifted far from its purely scientific origins. He presents substantial and meaningful information and anecdotes about the academic arms race, the drive for comprehensive funding and the complexities and perils of corporate funding of research. He guides us through the negative effects that can result throughout the academic system for a large-scale, high-impact enterprise that is highly competitive at individual and institutional levels.

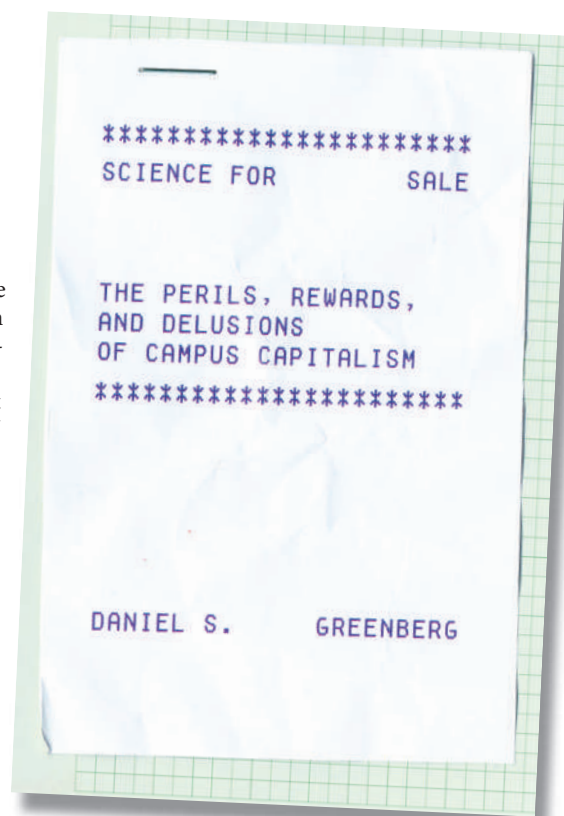
Academics would admit that Greenberg's concerns about conflicts of interest and the challenges of funding are legitimate and that these need to be managed consciously and meaningfully. But his assumptions and arguments about the corruption of academic purpose and of its independence from

corporate and government funding are oversimplified. He presents scenarios in which the actors involved, be they academic scientists, university administrators or corporate technologists, do not seem to live up to his ideals in terms of their scientific comportment; his analyses aren't so much incorrect as lacking in detail and historical context.

From its beginnings, US academic science has focused on practical outcomes and has been linked heavily to industry. Take the Silliman family of Connecticut. Benjamin Silliman, a Yale alumnus and son of the revolutionary war hero General Gold Selleck Silliman, switched professions from law to medicine and in 1804 taught the first academic science class in the United States. He was a hybrid chemist, geologist and science-society organizer who, between 1804 and 1857, spent much of his time on practical mineralogy and consulting for industry and the government. Silliman's son, also Benjamin, became a Yale chemist and one of the most influential academic scientists in US history. He wrote a US\$526 report for industry on how to distil fractionated oil, which ultimately led to the emergence of the oil industry in Pennsylvania and provided the chemical source for the Standard Oil Company. These path-breaking academics helped shape the structure of America's first school of science (Sheffield Scientific School at Yale) and furthered the US tradition of practical science.

This account and numerous others demonstrate the emergence of US universities through donation, state constitution, church, land-grant or powerful local business and commercial forces. *Science for Sale's* portrayal of academic science as previously unencumbered by market forces, greed or evil is not the whole story.

The theory Greenberg presents is desirable. In reality, science and its environment represent a much more complicated force. Academic scientists work to advance humanity, sometimes coming up with innovations that have the potential for good and bad outcomes. One of Greenberg's case studies of corporate influence on the development trajectory of the chemotherapy drug Taxol seems like child's



play when compared with government influence on the production of new technologies with a high potential for social and economic disruption. Even science at its most pure can be affected by a variety of factors, sometimes with unintended consequences. The point is, all influences are at least equal, all outcomes can cut both ways and all resources carry with them a *quid pro quo*.

Greenberg outlines the myriad pitfalls and pitfalls faced by academic science, but does not analyse how to address them. US research universities exist inside one of the bastions of capitalism and are organized to compete with each other for students, faculty, resources and recognition in this complex sociocultural system. Unlike much of the rest of the world, they are not managed by a central ministry. This setting has led institutions to take on many of the characteristics that concern Greenberg — size, scope, wealth, reliance on external resources, and so on. It is not a set of corporate or capitalistic forces that are driving US academia on the path of increased complexity and stress. It is the universities, through their creativity, connectivity, influence and power, that are driving other institutions or nation states to action with science-based competition. ■

Michael M. Crow is president of Arizona State University, PO Box 877705, Tempe, Arizona 85287-7705, USA.

UNIV. CHICAGO PRESS

Atlas on our shoulders

The Body has a Mind of its Own: How Body Maps in Your Brain Help You Do (Almost) Everything Better

by Sandra Blakeslee and Matthew Blakeslee

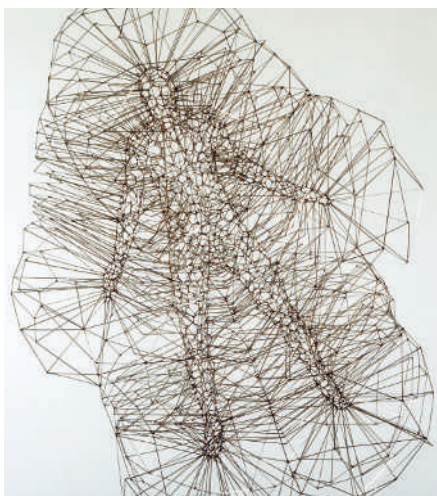
Random House: 2007. 240 pp. \$24.95

Edvard I. Moser

Like an atlas, the brain contains maps of the internal and external world, each for a distinct purpose. These maps faithfully inform the brain about the structure of its inputs. The body surface, for example, is mapped in terms of its spatial organization, with the same neural arrangement flashed through successive levels of processing — from the sensory receptors in the periphery to the thalamus and cortex in the brain. Meticulous mapping also takes into account the hat on your head and the golf club in your hand, and separate systems encode the space through which your body moves. As V. S. Ramachandran has remarked: God must be a cartographer.

In *The Body has a Mind of its Own*, Sandra and Matthew Blakeslee — who represent the third and fourth generations in a family of science writers — use this emerging knowledge about brain maps as a springboard for presenting some of the most exciting discoveries in neuroscience. The unifying theme is the idea that the way our body is mapped by neural circuits in the brain can account for a range of our experiences and perceptions. Using a readable and inspiring format, the authors showcase new and classic research on neural representations, without compromising accuracy.

The stories are refreshingly up-to-date, covering, for example, the emerging appreciation of the dynamic borders between maps for body and external space, and new insights



Sculptor Antony Gormley also explores how the body relates to surrounding space.

into how our own location in space is mapped by specialized cells called grid and place cells. Anecdotes and ideas from sister disciplines, including neurology, psychiatry and cultural anthropology, mix comfortably with laboratory observations. New discoveries titillate our curiosity, explaining common phenomena such as yo-yo dieting and contagious yawning as well as some more bizarre neurological abnormalities such as alien-hand syndrome and supernumerary-limb perception. Also covered are why you cannot tickle yourself, why some people have 'out-of-body' experiences, and why babies in Mali walk earlier than those anywhere else in the world.

The chapter on mirror neurons discusses their possible implications for social transmission and cultural learning. Mirror neurons fire when a monkey peels a banana, say, and when it observes another monkey peeling a banana.

They may therefore be part of the neural basis of intention, promoting learning by imitation. The authors explain how mirror neurons could participate in a wide range of primate brain functions, for example in shared perception and empathy, cultural transmission of knowledge, and language. At present we have few, if any, clues as to how mirror neurons compute or how they interact with other types of neuron, but the Blakeslees draw our attention to social neuroscience as an emerging discipline.

It is important to keep in mind that the map concept is not explanatory. We need to define what a map is to understand how perception and cognition are influenced by the spatial arrangement of neural representations. The classical maps of the sensory cortices are topographical, with neighbouring groups of neurons representing neighbouring parts of the sensory organ and thus the outside world. Other maps, such as those of the brain's spatial-representation system, are distributed with no striking correlation between firing properties and anatomical location. Both are maps in the sense that they link a feature of the external world with the activity of a neuron, but the functional implications of the map arrangements are likely to be vastly different. More complex wiring diagrams may be needed to maintain coherent localized representations in a distributed network than in a topographic system.

To the non-specialist, it may not matter whether a map is a map or not. *The Body has a Mind of its Own* is a thought-provoking book of wide appeal. It is a striking example of how complex issues in contemporary research can be presented to entertain everyone. ■

Edvard I. Moser is professor of neuroscience at the Kavli Institute for Systems Neuroscience and the Centre for the Biology of Memory, NTNU, Trondheim, Norway.

COURTESY OF THE ARTIST & J. JOPLING/WHITE CUBE

Ups and downs of a senator scientist

The Volterra Chronicles: The Life and Times of an Extraordinary Mathematician 1860-1940

by Judith R. Goodstein

American Mathematical Society/London Mathematical Society: 2007. 310 pp. \$59

Salvatore Coen

Vito Volterra was an extraordinary mathematician. He was also a pioneer of biological modelling — a research tool that has increased in popularity by leaps and bounds since the publication of his seminal 'predator-prey' model in 1926. Volterra lived at a time when Italy was undergoing bouts of political upheaval, as a child in the wake of the

Risorgimento — the country's unification in the 1860s — and then under Benito Mussolini's fascist regime. As a senator, Volterra withheld his support for the regime and was therefore expelled from the University of Rome in 1931 and from all Italian academies in 1934. *The Volterra Chronicles* is a rich biography set amid this turmoil.

Author Judith Goodstein captures the human, political and social environment of the age. The story of Volterra's life stems from his very close Italian Jewish family, with their strong patriotic sentiments. He was born in Ancona in May 1860, when the town belonged to the Church, and grew up in a country still deeply affected by the ideals of

the *Risorgimento*. He graduated in physics from the University of Pisa and became a professor there when he was only 23 years old. Later he moved to the University of Turin, then to the University of Rome, where he married his young cousin Virginia Almagià.

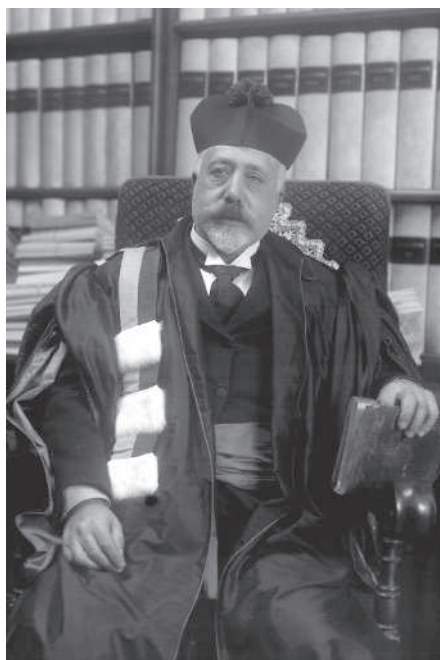
The Volterra Chronicles describes the early research done by the young Volterra and his deep and precocious interest in mathematics and its applications. His papers on real and complex analysis, on mathematical physics, hydrodynamics, electrostatics, elasticity, differential, integral and integro-differential equations are still classics. In addition, he is considered a father of functional analysis, one of the main branches of contemporary

mathematics. But he never lost contact with the experimental sciences, and the final chapter of the book covers his ground-breaking incursions into biomathematics.

Volterra was a key participant in scientific organizations in Italy and worldwide. He was president or vice-president of important research institutes including the International Bureau of Weights and Measures (1921–40) and the International Research Council, and founder of, among other things, the Italian National Research Council and the Italian Society for the Progress of Science.

His family has generously donated several archives to the Accademia dei Lincei in Rome, which are now fully catalogued. This biographical material has been invaluable for scholars, and important contributions about Volterra have been published during the past decade or so — for example, by Giorgio Israel and Ana Millán Gasca in 2002 and Rita Levi-Montalcini in 1996.

The Volterra Chronicles is different from these other works in that it is a dedicated biography of Volterra, based on original documents, which invites us into the mathematician's intimate circle. Short passages translated from his correspondence provide insight into his intense relationships with the most prestigious mathematicians of his time, which were always grounded in mutual scientific and personal respect. The book reveals his knack for spotting talent in young scientists, including Tullio Levi-Civita, André Weil,



Vito Volterra modelled predator-prey dynamics.

Griffith Evans, Marcel Brelot, Enrico Fermi and Joseph Pérès, whom he helped in various ways or with whom he collaborated. The characters of his family and friends, including Carlo Somigliana and the physicist Antonio Ròiti, and of his teachers Ulisse Dini and Enrico Betti are perceptively drawn.

Volterra was resolutely in favour of mili-

tary intervention by Italy during the First World War to fight alongside France, Britain and Russia. He volunteered at the age of 55 when he was a senator and a member of many illustrious academies. During the conflict, he was involved in military action and scientific research, and cooperated closely with scientists in allied countries, particularly France.

He was a fierce opponent of fascism. In 1924, still a senator, he cast a no-confidence vote in Mussolini as head of the government and in 1931, he refused to sign a loyalty oath professing allegiance and devotion to king, country and the Fascist regime. This refusal cost him his university post and academy memberships, as well as the support of most of his colleagues. He also faced persecution under racial laws issued in 1938. But the regime allowed him to keep his passport and he continued to travel until his death in October 1940, four months after Italy joined the Second World War.

One of three appendices (the others being translations of his most important speeches) to *The Volterra Chronicles* reproduces a touching and sensitive obituary written by the distinguished English mathematician Edmund Whittaker. It is a fitting end to a biography of an outstanding scientist, which is remarkable for its accuracy and scholarly rigour, and — unlike its predecessors — for its potential to engage a wide readership. ■

Salvatore Coen is professor of geometry in the Department of Mathematics at the University of Bologna, 40126 Bologna, Italy.

Oceanography fathomed

To Follow the Water: Exploring the Ocean to Discover Climate

by Dallas Murphy

Basic Books: 2007. 288 pp. \$26

Arnold Gordon

The ocean covers nearly 71% of our planet, providing food, minerals, energy, transport, even recreation. It is also crucial in governing our climate system. Indeed the 29% of solid land, which depends on the sustained health of the entire environmental system, is arguably itself anchored in the ocean.

In *To follow the Water*, Dallas Murphy gives the general reader and the specialist an ocean view beyond the confines of the laboratory or computer model. He rolls from topic to topic in a readable and scientifically astute narrative that touches on the oceans' many aspects — from their exploration over the centuries to their current impact on climate — and their study as a quantitative science today.

Fundamental discoveries await oceanographers, but observations are difficult; the oceans are vast and controlled experiments are impossible. I entered physical oceanography decades



Remotely operated vehicles, controlled from aboard ships, measure the ocean's fluctuations.

ago out of curiosity, coupled with the excitement of going to sea. The field has expanded hugely since then: the research is still fun, but it is now more essential than ever to get it right.

In the early days of exploration, European nations competed to find sea routes to Asia around the southern rim of Africa, or to the great unknown to the west. The southern route succeeded; the western route led to gold, to the great misfortune of those who had arrived there earlier. Even then, some explorers were intrigued by ocean currents as a means to speed up their journey, and by the origin of the cool, tropical subsurface water in which they cooled off.

Gradually these explorers morphed into scientists, although sea-going oceanographers still retain an element of the explorer. We are now armed with more sophisticated tools to measure the ocean's fluctuations — from aboard ships, from instruments moored or floating freely in the ocean, and from satellites peering down from above. Considering salt water's high density — granite is only 2.5 times more dense — and its high heat capacity and latent heat, it's no wonder the vast ocean mass has a lock on the climate. Unfortunately, our inability to simulate properly the ocean processes that transfer heat, water and dissolved chemicals undermines the validity of our global climate models. And there is that annoying dab of carbon dioxide that we are so unwisely burning into the atmosphere, half of which ends up acidifying the ocean.

Oceanographers spend many weeks aboard ship, enduring the rolling and pitching that strains brain and belly. Murphy discovered this uncomfortable truth as a technician aboard the Woods Hole Oceanographic Institution's research vessel *Oceanus* and on two cruises aboard the US National Oceanic and Atmospheric Administration's *Ronald H. Brown*. He learned that big science, with its multifaceted programmes involving numerous institutions and countries, boils down to a small band of oceanographers tending to arrays of instruments over 24-hour watch schedules.

Murphy focuses on the North Atlantic meridional overturning circulation (MOC) — of which the Gulf Stream is part — only briefly mentioning other areas of the ocean. The MOC causes warm waters from the top kilometre of the ocean to spread into the northern North Atlantic, where they cool and sink to return southwards; in the process, western Europe is warmed. Fear of a breakdown in the MOC in a warming climate is what motivates much of the big science of the North Atlantic. Past climate patterns suggest that the MOC may slow during cool phases, not warm ones, but cause-and-effect issues need to be sorted out and a more global prospective is needed. Will global climate warming shut down the MOC and induce a cooling of northern Europe? I doubt it, but then the rapidly forced, anthropogenic modification of climate may not follow the same course as past natural changes.

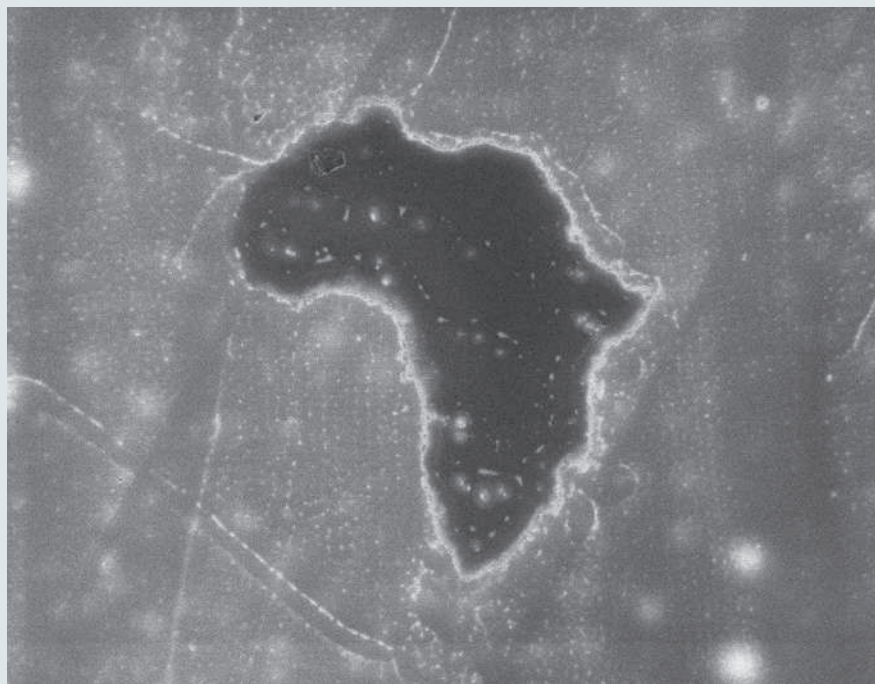
Murphy effectively points out that although the ocean sciences are now more important than ever, funding is so inadequate that very few worthwhile projects receive support. Chasing grant money reduces researchers' productive time, and delays advances in understanding the sea; and the clock is ticking. Meanwhile, the public needs to know about the oceans, but oceanographers remain wary

of media misrepresentation of their work. Meticulously following the waters of the Gulf Stream into the blue beyond, Murphy's book gets it right. ■

Arnold L. Gordon is professor of oceanography and associate director of the Division of Ocean and Climate Physics at the Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964-8000, USA.

EXHIBITION

The invisible continent



A tiny map of Africa, visible only with a microscope, represents a global paradox.

Emiliano Feresin

Satellite pictures of Earth have become too familiar to inspire awe. But two Turin-based artists have been prompted by the large-scale view of our planet — and by that other world, at the molecular level, that nanotechnologists can now reveal and manipulate.

Alessandro Scali from Turin, and Robin Goode, a South African, conceived *Actual Size*, the smallest ever map of Africa. The artwork is a tiny sliver of silicon just a thousand or so atoms wide, made with atomic force microscopy, and visible only with the same technique. It is one of the 'nanoartworks' to be shown in the duo's first exhibition next month, as part of the Bergamo science festival in Italy.

"We know from books

that Africa is the cradle of humanity, and we can see from maps and the beautiful satellite images that Africa is vast and is at the centre of the world," says Scali. "But we also know that Africa is poor, exploited, neglected." The artists wanted to represent the paradox of Africa being geographically and anthropologically central to our world, yet unrecognized, unexplored, even 'invisible'.

Scali and Goode brought their concept to Fabrizio Pirri, a physicist at Turin Polytechnic University. Pirri's group turned its atomic force microscope — usually used for biomedical engineering — to a modern form of lithography. Through electrical interaction, the microscope's 10-nanometre-wide tip oxidized individual

atoms of the silicon to trace the shape of Africa. They made dozens of copies before obtaining the unique piece that satisfied the artists. For the exhibition, Scali and Goode have made poster-size copies of their work, which could otherwise only be seen in the lab.

A second, larger version of *Actual Size*, as big as the point of a needle, will also be on show at Bergamo, and visible to visitors through a microscope. In this case, the scientists used a laser — which removes atoms through sublimation — to etch the map onto silicon. ■

Emiliano Feresin is an intern in *Nature's* Munich office.

The exhibition Nan°art runs 2-21 October in Bergamo, Italy.

COURTESY OF PHYSICS DEPT, TURIN POLYTECHNIC UNIV.

NEWS & VIEWS

ENVIRONMENTAL SCIENCE

Forests and floods

William F. Laurance

One strategy for reducing catastrophic floods invokes protecting or regenerating forests that lie upstream of a threatened region. Support for this approach comes from a global-scale analysis of flood risk.

Floods can be devastating, especially in developing countries, which are already struggling to improve the welfare of their people. Floods in such countries during the 1990s took almost 100,000 lives, displaced more than 300 million people and caused more than US\$1 trillion in damage to properties^{1,2}. To mitigate flooding, several nations are investing in projects to restrict the loss of forests or to reforest cleared lands that lie upstream of a potential flood region. But these strategies are controversial. Some argue that the link between forests and flooding is too weak to justify the often large outlays needed to retain or recover forests^{2,3}. In the journal *Global Change Biology*¹, however, Bradshaw *et al.* provide strong correlative evidence that native forests do reduce the frequency and severity of floods in developing nations.

Forests are thought to reduce flooding⁴ by acting as sponges — that is, they trap water during heavy rainfall, then release it slowly into streams, which lessens the severity of floods and maintains stream flows during dry periods. Forests also increase the permeability of the soil and emit water vapour into the atmosphere through evaporation and transpiration, further reducing the run-off of rainwater⁵. For these reasons, a nation such as Costa Rica, which places high value on natural-ecosystem services, and those such as China, India, Nepal and Bangladesh, which have been plagued by devastating floods (Fig. 1), have invested heavily in forest protection or reforestation.

Yet this approach is contentious. Most notably, a serious counterargument came from an influential report² published in 2005 by the Food and Agricultural Organization of the United Nations (FAO) and the Center for International Forestry Research (CIFOR) in Bogor, Indonesia. The analysis concluded that the evidence that forests reduce flooding is weak, especially for the largest and most devastating floods. The report suggested that retaining or regenerating large forest areas was an economically dubious strategy for developing nations, at least from a flood-reduction perspective.

Bradshaw *et al.*¹ step into the fray with a global-scale empirical analysis of the association between forests and flooding; earlier efforts had



R. BECK/AFP/GETTY

Figure 1 | Jiujiang, China, 12 August 1998. Local dyke defences in the Jiujiang region were overcome during a series of floods caused by relentless rain in the upper catchment of the Yangtze river.

invariably examined much smaller catchment areas or regional scales. The authors evaluated flood frequency, duration and damage (the number of people killed and displaced, and estimated economic losses) from 1990 to 2000 in 56 developing countries in Africa, Asia and Latin America. Using a battery of statistical models, they first controlled for confounding factors such as country size, rainfall, topography, degraded land and soil-moisture regime, then evaluated the residual associations between native-forest cover and flooding.

Their simplest and most realistic models account for more than 65% of the variability in flood frequency and duration, an impressive figure for such a vast, heterogeneous cross-section of Earth. Native-forest cover and the rate of forest loss account for 14% of this variation, on average. Associations between forest cover and the damage caused by floods were weaker but still evident. Statistical simulations

suggested that arbitrarily removing a tenth of the remaining native forest would increase the frequency of floods by 4–28%, and lengthen their duration by 4–8%. The authors emphasize that assessment of flood risk is complex and fraught with confounding factors. But they say that their study shows the need for large-scale forest protection and more reforestation to help reduce the frequency and severity of floods.

In my view, Bradshaw and colleagues' paper¹ will turn out to be a landmark study. But at least one point will provoke debate — their decision to exclude floods driven by extreme events, such as cyclones and typhoons, which they suggest can cause flooding “independently of landscape characteristics” such as forest cover. Their rationale is that massive storms can dump vast amounts of rainfall in just a few days. Even if upstream forests were intact, almost nothing could prevent downstream lowland areas from flooding.

The apparent inability of forests to buffer against extreme floods is one of the main arguments of the FAO–CIFOR team², who assert that expensive efforts to reforest steep catchments in densely populated, monsoonal countries such as Nepal, India and Bangladesh have been largely ineffective. Rather, they argue, investments would be better aimed at other measures, such as discouraging human settlement in flood plains, which can be devastated by flooding during monsoons. Bradshaw *et al.* side-step the issue of extreme flood events. But if nations are to be convinced that forests are valuable for reducing flooding, it will be necessary to show that they are of at least some benefit during the most catastrophic floods. In addition, the broad-brush approach of Bradshaw *et al.* cannot encompass all situations. Because local conditions vary greatly, forests will surely be better at reducing floods in some areas than in others.

Nonetheless, the study's ambitious geographical scope and statistical rigour make it persuasive. The findings add to other evidence that large expanses of native forest can have major benefits not only for reducing floods, but also for limiting wild fires, conserving biodiversity, and slowing regional and global climate change^{4–6}. Tropical forests, in particular, are crucial for combating global warming, because of their high capacity to store carbon and their ability to promote sunlight-reflecting clouds via large-scale evapotranspiration⁷. Such features are key reasons why preserving and restoring tropical forests could be a better strategy for mitigating the effects of carbon dioxide than dramatically expanding global biofuel production⁸.

Flood reduction is one area of environmental science in which the latest research is quickly invoked to justify land-management strategies, with all the pungent realities and controversies this implies. In a world facing potentially serious shifts in rainfall, water run-off⁹ and storm regimes¹⁰ from climate change, and where burgeoning populations are increasingly forced to live in flood-prone areas, the stakes could not be higher.

William F. Laurance is at the Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancón, Panama.
e-mail: laurancew@si.edu

IMMUNOLOGY

Changed destiny

Huafeng Xie and Stuart H. Orkin

For practical and ethical reasons, researchers are on the lookout for ways to reprogramme one mature cell type into another. In one case, this might be as easy as switching off a single gene.

The norm is for most undifferentiated, progenitor cells to irreversibly mature into one particular type of cell. But some cells, even after commitment to a specific fate, can acquire properties of an entirely different cell through a process called cellular reprogramming.

An *in vitro* example of reprogramming is the conversion of skin cells into cells with characteristics of embryonic stem cells by adding a small set of nuclear factors to them^{1–3}. The molecular details of how reprogramming is carried out are not known. In this issue, Cobaleda *et al.*⁴ (page 473) shed light on this mystery by showing that, within the blood system, mature B cells can be converted to functional T cells, and reprogramming is achieved by B cells taking a step backwards to assume a more immature state.

As progenitor blood cells embark on their maturational path, the developmental options available to them become progressively more limited⁵. Lineage commitment is principally controlled by transcription factors that act as master regulators. Examples of these include PU.1 and C/EBP α , which control the commitment of myeloid cells; GATA-1 and FOG-1, which are involved in the development of erythrocytes and megakaryocytes; and Notch-1, GATA-3 and Pax5, which mediate T- and B-cell development.

Conventional wisdom holds that once a cell is committed to a given lineage through the concerted actions of transcription factors and epigenetic modifications to chromatin (complexes of DNA and histone proteins), its fate is sealed. In recent years, however, the unidirectional nature of differentiation has been questioned after reports that cells of the blood lineage are more 'plastic' than previously thought — that is, through various manipulations, they can be reprogrammed to a different lineage⁶. The fate of blood cells can be altered by introducing specific transcription factors not normally expressed in them. For example, introducing C/EBP α into B or T cells converts them into functional macrophages^{7,8}. GATA-1 can also reprogramme common B and T progenitor cells into megakaryocytic/erythroid cells⁹.

In principle, cellular reprogramming may proceed in one of three ways: by de-differentiation to a progenitor stage, followed by differentiation along a different route; by direct transdifferentiation from one lineage to another; or through a process involving both de-differentiation and transdifferentiation (Fig. 1).

In transdifferentiation, a new developmental

programme is established before the original one is terminated. And so cells pass through a stage in which they show characteristics of two lineages. For example, during the process of C/EBP α -induced reprogramming of B and T cells to macrophages, cells pass through an intermediate stage^{7,8}. In a subset of these intermediate cells, macrophage characteristics are dominant and switch off mechanisms that maintain B- and T-cell characteristics.

Expression of a master regulator is not the only way a cell's fate can be altered; loss of these regulators may also result in cellular plasticity. Whether this takes place through a trans-differentiation or a de-differentiation route is not known. For example, loss of Pax5 halts development of the B-cell lineage at the earliest (progenitor B cell) stage^{10,11}. In cell culture or after transplantation into irradiated hosts (whose normal blood-cell progenitors are destroyed), Pax5-deficient B-cell progenitors can mature into all blood lineages such as myeloid and T cells — except B cells^{10,11}. In mature B cells, inactivation of the *Pax5* gene results in B-cell loss through an unknown mechanism¹².

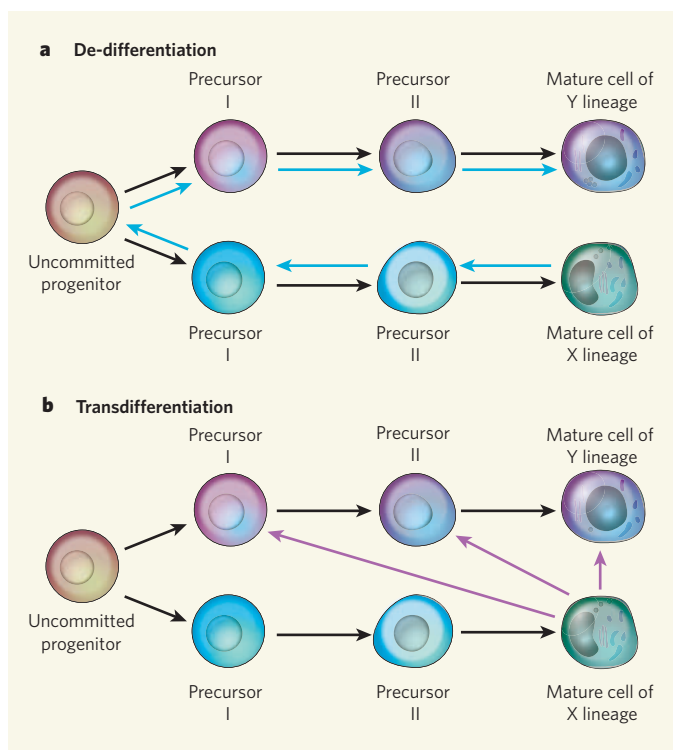
Cobaleda *et al.*⁴ now show that, in mice, loss of *Pax5* in mature B cells leads to a type of transplantable cancer known as malignant lymphoma in which blood cells show characteristics of progenitor cells. However, mature B cells are still present in the secondary lymphoid organs (such as lymph nodes and spleen) of these animals. When transplanted into mice that lack all lymphocytes (B cells, T cells and natural killer cells), B cells that lack *Pax5* efficiently generate T cells by first reverting to pro-B cells. The authors confirm the 'durability' of the reprogrammed B cells by showing that, when purified pro-B cells from recipient mice are transplanted into secondary recipient mice, they can still support T-cell development. The newly acquired T cells in the recipient mice were found at all stages of development, with the overall frequency of each subpopulation showing its usual distribution. Importantly, the reprogrammed subpopulation of mature T cells was functional, responding to antigen stimulation.

The work of Cobaleda and colleagues clearly indicates that, in the context of Pax5-deficient B cells, cellular reprogramming takes place through de-differentiation. It also raises many questions. For example, is de-differentiation an essential step in cellular reprogramming? If so, must cells return to an uncommitted, 'ground' state, or is direct reprogramming

- Bradshaw, C. J. A., Sodhi, N. S., Peh, K. S.-H. & Brook, B. W. *Global Change Biol.* doi:10.1111/j.1365-2486.2007.01446.x (2007).
- FAO–CIFOR *Forests and Floods: Drowning in Fiction or Thriving on Facts?* (FAO–CIFOR, Bangkok, 2005).
- Calder, I. R. & Aylward, B. *Water Int.* **31**, 87–99 (2006).
- Clark, C. *Environ. Conserv.* **14**, 67–69 (1987).
- Bruijnzeel, L. A. *Hydrology of Moist Tropical Forest and Effects of Conversion: A State of Knowledge Review* (UNESCO, Paris, 1991).
- Laurance, W. F. & Peres, C. A. (eds) *Emerging Threats to Tropical Forests* (Univ. Chicago Press, 2006).
- Bala, G. *et al. Proc. Natl Acad. Sci. USA* **104**, 6550–6555 (2007).
- Righelato, R. & Spracklen, D. V. *Science* **317**, 902 (2007).
- Betts, R. A. *et al. Nature* **448**, 1037–1041 (2007).
- Webster, P. J., Holland, G. J., Curry, J. A. & Chang, H.-R. *Science* **309**, 1844–1846 (2005).

Figure 1 | Ways of cellular reprogramming.

a, It is possible that a mature cell lineage (X) de-differentiates towards an uncommitted progenitor from where it would follow a different developmental path to eventually become — through a different set of precursor stages — a mature cell of another lineage (Y). Cobaleda *et al.*⁴ find that this is the process through which Pax5-deficient mature B cells give rise to mature T cells. **b**, Alternatively, mature X cells might directly transdifferentiate into precursor or mature Y cells. A third route might involve simultaneous de-differentiation and transdifferentiation.



from intermediates possible? Cellular reprogramming, in contrast to nuclear reprogramming in eggs, which occurs within hours, is inefficient and takes considerable time (several days to weeks). So what is the sequence of molecular events leading to reprogramming? What are the rate-limiting steps?

Do T cells derived from de-differentiated B cells carry epigenetic marks of their B-cell heritage or are these replaced by marks characteristic of T cells? Can mature B-cell-derived T cells return to the B lineage upon expression of Pax5, or will they form T-cell lymphoma cancers, which occur

when Pax5 is uncharacteristically expressed in normal T cells?

Another question is whether there are mechanisms used both in B-cell conversion to other lineages and in the reprogramming of skin cells to an embryonic state, as seen in earlier studies^{1–3}. Comparative analysis of reprogramming in different cell systems should further our understanding of the molecular events crucial to the process, and suggest strategies for enhancing the efficiency of converting one lineage to another. Such insights will, in turn, make the alteration of cell fates using modulation of gene expression and the generation of a specific cell population possible, which is a primary goal of regenerative medicine. ■

Huafeng Xie and Stuart H. Orkin are in the Department of Pediatric Oncology, Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA.
e-mail: stuart_orkin@dfci.harvard.edu

- Okita, K., Ichisaka, T. & Yamanaka, S. *Nature* **448**, 313–317 (2007).
- Wernig, M. *et al.* *Nature* **448**, 318–324 (2007).
- Maherali, N. *et al.* *Cell Stem Cell* **1**, 55–70 (2007).
- Cobaleda, C., Jochum, W. & Busslinger, M. *Nature* **449**, 473–477 (2007).
- Orkin, S. H. *Nature Rev. Genet.* **1**, 57–64 (2000).
- Graf, T. *Blood* **99**, 3089–3101 (2002).
- Laiosa, C. V. *et al.* *Immunity* **25**, 731–744 (2006).
- Xie, H., Ye, M., Feng, R. & Graf, T. *Cell* **117**, 663–676 (2004).
- Iwasaki, H. *et al.* *Immunity* **19**, 451–462 (2003).
- Rolink, A. G., Nutt, S. L., Melchers, F. & Busslinger, M. *Nature* **401**, 603–606 (1999).
- Nutt, S. L., Heavey, B., Rolink, A. G. & Busslinger, M. *Nature* **401**, 556–562 (1999).
- Horchner, M., Souabni, A. & Busslinger, M. *Immunity* **14**, 779–790 (2001).

NUCLEAR PHYSICS

A non-disappearing magic trick

Ragnar Bengtsson and Peter Möller

Well-established models of nuclei describe properties such as shells and magic numbers. But how do these predictions stand up to scrutiny for exotic, unstable nuclei? Pretty well, according to the latest study.

Of the very many atomic isotopes that could possibly exist in nature, only a few hundred are actually stable — just a handful, on average, for each chemical element. No stable isotope exists at all beyond bismuth (proton number $Z = 83$), although some isotopes of thorium ($Z = 90$) and uranium ($Z = 92$), with half-lives in the billions of years, come close. The rest are unstable. But even though these elements might exist for only a fraction of a second before decaying to something more stable (often through a sequence of reactions), they are not uninteresting: unstable isotopes pop up not just in the ashes of man-made nuclear reactors, but also in stars including the Sun, and in exploding supernovae. Writing in *Physical Review Letters*, Seweryniak *et al.*¹ report a breakthrough in understanding the proper-

ties of one group of highly unstable nuclei near the tin isotope ^{100}Sn . This isotope has both a Z of 50 and a neutron number N of 50.

This conjunction is significant. The context is nuclear ‘magic numbers’ — values of Z and N at which many nuclear properties, which elsewhere vary smoothly, change suddenly and drastically. These jumps are associated with the largest deviations in experimental nuclear binding energies — the energy needed to break up a nucleus into individual, well-separated protons and neutrons (nucleons) — from values predicted by the simple, but remarkably accurate, liquid-drop model. This model predicts nuclear binding energies of about 8 megaelectronvolts (MeV) per nucleon, but with significant variations throughout the nuclear chart.

For a heavy nucleus, the total binding energy

approaches 2,000 MeV, and the liquid-drop model predicts the values to an average accuracy of about 3 MeV. But at the magic numbers (which are $Z = 2, 8, 20, 28, 50$ and 82 , and $N = 2, 8, 20, 28, 50, 82$ and 126) the observed total binding energies are up to 10 MeV higher than predicted by the liquid-drop model. These nuclei are thus significantly more stable than the norm.

In 1950, Maria Goeppert-Mayer and Hans Jensen² showed that these magic numbers could be extracted from the solutions to an appropriate microscopic, quantum-mechanical Schrödinger equation. The form of this equation assumed a system of nucleons that was governed by a ‘single-particle potential’ coupled with a strong spin–orbit force. The solutions to this equation predicted many other nuclear properties, in particular the observed energies of excited states and their quantum numbers. In this, the first nuclear-shell model, magic numbers represented points of particular stability that were obtained when a shell containing a set number of nucleons became full. This outstanding theoretical achievement won Goeppert-Mayer and Jensen a share of the 1963 Nobel Prize in Physics.

Because both its Z and its N are 50, ^{100}Sn is in fact doubly magic. In this respect, it is in rather

exclusive company: only eight other doubly magic nuclei are known to exist. Five of the nine — ^4He , ^{16}O , ^{40}Ca , ^{48}Ca and ^{208}Pb — are stable. Spectroscopic data from these nuclei were central to the development of the nuclear-shell model. But being doubly magic doesn't automatically confer stability. Far from the central axis of the long, thin strip of the nuclide chart (Fig. 1), nuclei are increasingly unstable, and tend, through β -decay, to move towards the central line (the 'line of β -stability'). Being doubly magic does, however, make a nucleus longer-lived than the nuclei around it. Of the remaining four doubly magic nuclei, two are nickel isotopes, ^{56}Ni and ^{78}Ni , which have β -decay half-lives of 6 days and 110 milliseconds, respectively. The other two are tin isotopes: ^{132}Sn , with a half-life of 40 seconds; and ^{100}Sn , studied by Seweryniak and colleagues, with a half-life of 1 second.

Of these isotopes, ^{56}Ni is only two neutrons away from its nearest stable isotope, ^{58}Ni , and ^{132}Sn is four protons away from its nearest stable isotope, ^{136}Xe . In both cases, single-particle energy levels (which can be occupied by a nucleon excited to a higher energy state) have been observed experimentally for several decades. For the truly exotic nuclei, ^{100}Sn (nearest stable isotope 8 protons away) and ^{78}Ni (8 neutrons away), no spectroscopic data had been available. These nuclei are hard to produce, and were first detected only in the 1990s. But as Seweryniak *et al.*¹ point out, better assessment of the characteristics of these systems would be very useful, in particular to test whether predictions of magic numbers and the order of single-particle levels obtained in models with the standard spin-orbit force² are still valid this far from stability.

Some idea of the significance of these nuclei can be gained from Figure 1, which depicts the deviation of nuclear binding energies from those predicted by the macroscopic liquid-drop model. Here, the four doubly magic nuclei heavier than ^{56}Ni rise as isolated mountain peaks above the vast expanse of the liquid-drop-model landscape. The isotope ^{100}Sn lies just inside the proton drip line — the boundary beyond which a nucleus decays more or less instantaneously by emitting a proton — and is the heaviest-known doubly magic nucleus that has an equal number of protons and neutrons.

Across the valley of β -stability, ^{78}Ni lies on the other, neutron-rich side. It is especially interesting as it is the only doubly magic nucleus that lies at a point where it is susceptible to so-called r -processes. These are a succession of nucleosynthesis reactions, consisting of repeated steps of neutron capture and β -decay that plot a zig-zag course through the nuclide chart. These

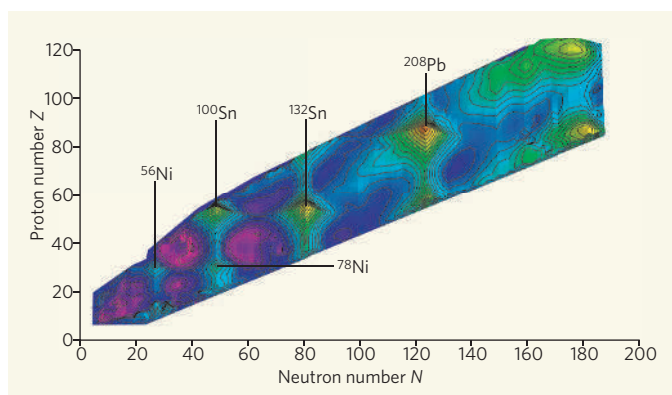


Figure 1 | The microscopic nuclear landscape. This depiction of 'microscopic corrections' across the nuclide chart shows the nuclear binding energy of individual nuclides calculated in a modern model that combines a macroscopic approach (the liquid-drop model) with microscopic nuclear-shell models originally proposed by Goeppert-Mayer and Jensen². To allow variations in the microscopic effects to be clearly visible — despite their importance to nuclear stability, they amount to a maximum of around 10 MeV relative to a total nuclear binding energy of up to 2,000 MeV in heavy systems — the binding energy given by the macroscopic model is not included. Great mountain peaks in the microscopic contributions manifest themselves at doubly magic proton and neutron numbers — among them ^{100}Sn , the subject of Seweryniak and colleagues' interest¹.

processes are thought to be responsible for creating most isotopes heavier than zinc ($Z = 30$) on the neutron-rich side of β -stability when they occur in core-collapse supernovae.

The single-neutron levels in the nuclear shell that is filled between the magic numbers $N = 50$ and $N = 82$ are known as hole states in ^{132}Sn , because they are revealed by removing a neutron. They are called particle states in ^{100}Sn , because a neutron must be added to reveal them. A reliable theoretical model should be able to describe how the energies of these states differ between these two nuclei. In the future, when spectroscopic data become available for ^{78}Ni , similar comparisons could be made between ^{100}Sn and ^{78}Ni , and between ^{78}Ni and ^{56}Ni . For a model to give reliable predictions of the properties of nuclei along the r -process path or in the region of superheavy elements around $Z = 114$, it should at least pass these tests.

As Seweryniak *et al.*¹ show, things are looking rather good for the models. First, most microscopic model calculations — including the original models of Goeppert-Mayer and Jensen² — that were published before any data were known for ^{100}Sn and adjacent nuclei, predicted the order of nuclear energy levels above $N = 50$ that these authors now find, which correspond to the ground state and first excited state of ^{101}Sn . Second, a mass excess of -56.8 MeV for ^{100}Sn , calculated in 1975³, is in excellent agreement with the experimental value of -56.78 ± 0.71 MeV. Results of several other microscopic mass calculations also agree with this mass measurement to within 1 MeV or so. The accuracy of these mass predictions indicates that the magnitudes of the magic level gaps at $N = Z = 50$ were also correctly predicted.

It is sometimes speculated that conventional magic numbers disappear far from stability. And yet where, as for ^{100}Sn , nuclear properties are found to be as predicted by a simple 50-year-old model², that charge can hardly be made to stick. But further data, particularly from spectroscopy, on the neutron-rich side near ^{78}Ni will be needed to deliver a final verdict.

Ragnar Bengtsson is in the Division of Mathematical Physics, Lund Institute of Technology, Lund University, Box 118, SE-22100 Lund, Sweden. Peter Möller is in the Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545-0001, USA. e-mail: moller@lanl.gov

1. Seweryniak, D. *et al.* *Phys. Rev. Lett.* **99**, 022504 (2007).
2. Mayer, M. G. & Jensen, J. H. D. *Elementary Theory of Nuclear Shell Structure* (Wiley, New York, 1955).
3. Seeger, P. A. & Howard, W. M. *Atom. Data Nucl. Data Tables* **17**, 428–430 (1975).

EVOLUTIONARY BIOLOGY

Development with a bite

P. David Polly

Predictions drawn from experiments on the growth of molar teeth in mice, and applied to a wide range of related species, provide a remarkable connection between developmental and evolutionary biology.

Palaeontologists know that mammals have lost and gained teeth in a regular evolutionary pattern: incisors from the back of the incisor series, premolars from the front of the premolar series, and molars from the back of the molar series (Fig. 1, overleaf). They also know that teeth in those series typically grade in size

in the same directions. And they know that it is rare that a species evolves more molars than the ancestral number of four in marsupial mammals or three in placental mammals. On page 427 of this issue, Kavanagh *et al.* describe a simple developmental cascade system that is eerily accurate in predicting the number and



50 YEARS AGO

Modern technology is confronting us with an exceedingly perplexing biological problem... It is the problem of how men and communities can adapt themselves to an environment which is changing with unprecedented speed [and] confronts teachers of all kinds at every level of education. One aspect of it — the higher education of technologists — is... specially important because technologists are now becoming the pacemakers for social change... The technologist is up to his neck in human problems whether he likes it or not. Take a simple example: the civil engineer who builds a road into a new territory in tropical Africa. He may assert that it is not his business to take into account the effect his road will have on primitive villages up-country... but he cannot afford to be utterly ignorant of the implications of his work.

From *Nature* 28 September 1957.

100 YEARS AGO

"Food inspection and adulteration" — [A] more drastic and far-reaching enactment is just now coming into force in the United States, and the working of one of its provisions in particular will be watched with much interest in this country. Its effect is to ensure that articles of food and drugs shall be labelled so as to show the purchaser, within limits, exactly what the articles are. The description must not be "false or misleading in any particular," whether as to composition, quality, origin, or what not. Thus an article must be stated on the label to be "prepared with glucose," "coloured with sulphate of copper," "dyed with aniline dye," or to be "composed of fragments and scraps from a mushroom cannery," and so on, as the case may be. Moreover, in the case of certain drugs — morphia, cocaine, chloral, chloroform, and others — the proportions must always be stated on the label.

From *Nature* 26 September 1907.

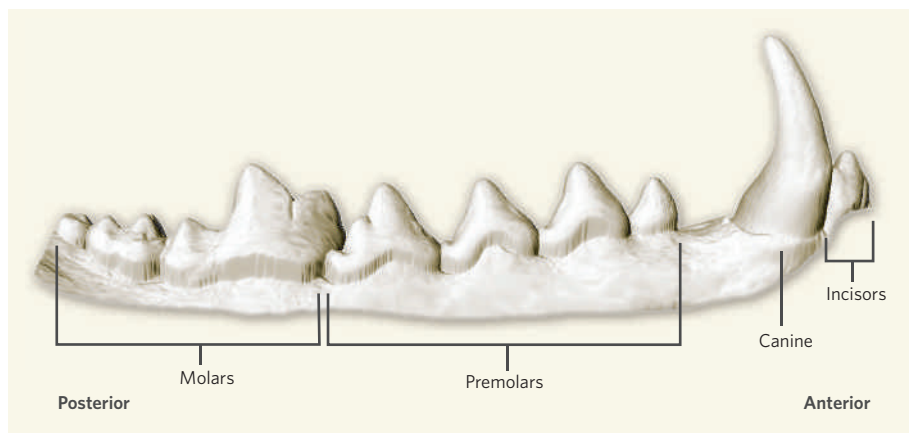


Figure 1 | Dentition of a placental mammal. This example — the lower teeth of a grey fox — shows the three-molar dental phenotype typical of placentals.

relative sizes of the molar teeth, so explaining how these seemingly arbitrary palaeontological observations are related to one another (K. D. Kavanagh, A. R. Evans & J. Jernvall *Nature* **449**, 427–432; 2007).

Embryonic molar teeth start as buds that spring from the dental lamina, a ribbon of epithelial tissue that runs parallel to the future tooth row. Buds initiate anterior-to-posterior, with the dental lamina growing in the same direction. Kavanagh *et al.* show experimentally that signalling molecules produced by developing mouse molars inhibit the development of subsequent buds. The balance between these inhibitors and activator molecules from the surrounding tissue determines when and if an additional molar will form. The higher the ratio of activator to inhibitor (a/i), the more rapidly molar buds will be added to the tooth row. And the more rapidly buds are added, the more there are and the bigger they get, meaning that a/i is a predictor of the relative sizes of the molar teeth (Fig. 2).

Using tooth buds growing in cell culture, Kavanagh and colleagues demonstrate these points by cutting the dental lamina behind the developing first molar. This interrupts the flow of inhibitors and allows the second molar to initiate earlier and grow to a larger size than normal. The inhibitor and activator molecules involved are probably the same as those active in the development of an individual tooth crown, such as Ectodin, Follistatin, Bmp3, Bmp4 and Activin βA .

If this developmental system is shared by all mammals, different dental phenotypes could be generated simply by varying the a/i ratio. Kavanagh *et al.* argue that the system has influenced the evolution of functional diversity in mammalian dentition. To test that possibility, they compile data on the proportional area of the molars of 29 species of murine rodents — close relatives of the mice in which the authors discovered the regulatory system.

The predictive mathematical model they derive from the developmental experiments explains nearly 75% of the diversity in molar proportions in these rodents. No species falls

far from the predicted proportions. The axis of dietary specialization parallels the axis of a/i , with herbivorous species at the activator heavy end of the developmental spectrum (where posterior molars are bigger than anterior ones) and animal-eating species at the inhibitor heavy end (where anterior molars are bigger). The authors convincingly argue that selection for diet may often act on the proportional expression of activators and inhibitors to produce a well-adapted dental phenotype.

The predictive power of their model is impressive, but will it hold for all mammals? From my further analyses, the answer is a qualified 'yes'. The results are shown in Figure 3, which depicts the 'morphological space' (morphospace) for different combinations of relative molar size. Nearly 70% of the variation from 35 additional species, representing 13 mammalian orders, is explained by Kavanagh and colleagues' model.

These new data probe the boundaries of their model by including species with phenotypes they did not test: marsupials (which typically have four molars); the bat-eared fox (an unusual

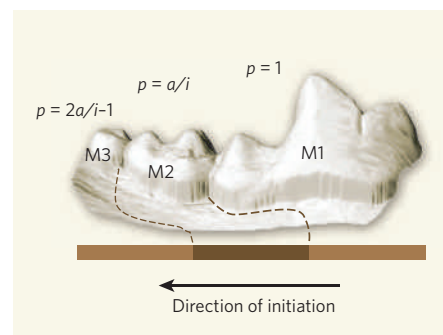


Figure 2 | Predictions of Kavanagh and colleagues' developmental model. The molar teeth — M1, M2 and M3 — develop from the front to the back. The size of the teeth are proportional (p) to the ratio of activator to inhibitor (a/i) molecules. Low a/i results in larger posterior molars, and high a/i results in larger anterior molars (like the ones shown here). Regardless of a/i , M2 will have an absolute size that is one-third of the combined size of all the molars (in species with three molars).

placental mammal with four molars); six extinct species (some as old as 55 million years); and species in which the first molar is much smaller than the second, which is in turn much smaller than the third ($M1 \ll M2 \ll M3$). All of these extreme phenotypes fall within the areas of morphospace consistent with Kavanagh and colleagues' model (the white areas of Fig. 3). Furthermore, plant-eating species tend to be in the high- a/i region, where $M1 < M2 < M3$ (upper white area), and animal-eating species tend to be in the low- a/i region, where $M1 > M2 > M3$ (lower white area), paralleling the pattern that Kavanagh and colleagues found in rodent species.

Areas of morphospace that are not consistent with their model are sparsely occupied (see the coloured areas in Fig. 3). Three species of bear fall deep in the region where $M1 < M2 > M3$, a pattern that Kavanagh *et al.* say requires early arrest of M3 development in addition to a decrease in inhibition. The horse falls (marginally) into the region where $M1 > M2 < M3$, which Kavanagh *et al.* say is the developmentally least likely phenotype. Members of the raccoon family, which have only two molars, are also an exception because their M2 is more than half the size of M1, larger than expected for animals that have lost M3.

Few developmental models derived from a single species are able to predict quantitative phenotypic variation across such huge evolutionary distances with such accuracy as does Kavanagh and colleagues' model. But tests need not end here. The model allows a/i to be estimated from the dental phenotypes of both extant and extinct species, and to be tested in the lab for extant species. The

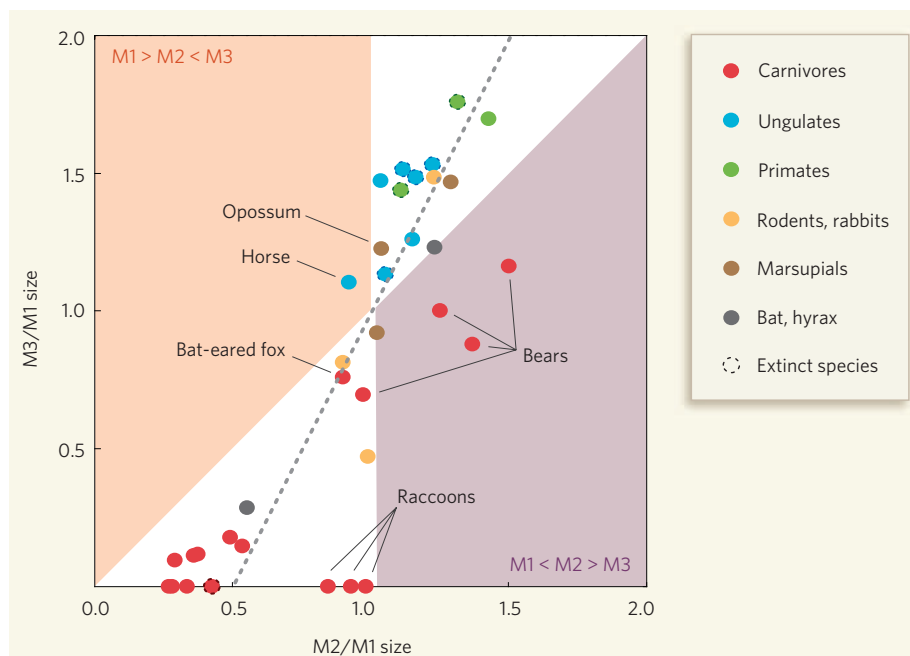


Figure 3 | Molar proportions in developmental 'morphospace'. Data from 35 species, compiled from my own measurements of specimens in the Indiana University Zooarchaeology Laboratory, show that Kavanagh and colleagues' model can largely account for patterns of molar-size proportions in mammals. The white region is consistent with their model; the broken line is the relationship they predicted. The bat-eared fox and opossum fit the model despite having four molars. Three species of bear do not fit the model (in that early arrest of M3 development is required in addition to a change in a/i); nor does the horse (though there is no obvious explanation for this). The raccoons have a proportionally larger M2 than expected for species without an M3.

study of signalling in raccoons or bears should reveal an interesting pattern of tooth-bud activation, inhibition and growth that deviates from Kavanagh and colleagues' findings in mice. ■

P. David Polly is in the Departments of Geological Science, and Biology and Anthropology, Indiana University, 1001 East 10th Street, Bloomington, Indiana 47405, USA.
e-mail: pdpolly@indiana.edu

QUANTUM PHYSICS

Qubits ride the photon bus

Antti O. Niskanen and Yasunobu Nakamura

Quantum mechanics using whole electrical circuits might seem a far-fetched idea. But make the circuits superconducting, and they can be used to send and collect single photons, rather like atoms do — only better.

The interaction of light and matter is all around us: we can see the objects that surround us only because their constituent atoms continuously emit and absorb electromagnetic radiation. Not only visible light, but everything from γ -rays through to radio waves, and even the alternating fields of power lines and the gigahertz signals inside a digital computer, are manifestations of fundamentally the same thing at different energy scales — the propagation of the discrete packets of electromagnetic energy known as photons.

In most situations, the presence or absence of a single photon does not make a noticeable difference to what happens. But this is not so in the nascent field of quantum computing, and

that explains the significance of two papers in this issue, by Sillanpää *et al.* (page 438)¹ and Majer *et al.* (page 443)². These authors describe experiments in which single photons transfer quantum information between relatively distant quantum bits (qubits) in a nanofabricated circuit held at a low temperature, so that the circuit loses its electrical resistance. Taken together with a closely related paper by Houck *et al.* in last week's issue³, which recounted how such superconducting circuits can be used to produce single photons on a chip, these papers represent confident steps towards the ultimate goal of a viable, large-scale quantum computer. But they also stand on their own as wonderful examples of how

science can mimic the ubiquitous natural interactions between atoms and light.

In principle, a serviceable quantum computer needs a collection of fully controllable qubits. A qubit can be represented by pretty much anything physical that can form a system of two states (representing 0 and 1) and behaves according to quantum theory. In practice, this means that the system must be as completely isolated from everything as possible. The most obvious choice for a qubit is one of nature's own-brand offerings — be they ions, atoms, molecules, electrons or photons. These microscopic objects can often be well isolated from their surroundings, but are quite difficult to control.

Despite their macroscopic size, specially designed electrical circuits provide a promising alternative basis for a quantum computer: their parameters can be adjusted, individual controls and measurements are relatively simple, and the system size can be scaled up. An essential ingredient for quantum computers is the ability to control the interactions between individual qubits for a 'universal entangling operation' to intertwine their quantum states. Qubits should be able to couple strongly, but they should also

GENOMICS

Vine work

The draft genome sequence of *Vitis vinifera*, the grapevine, described in this issue, provides plenty of scope for discussion over a glass of its fermented product. The sequence was published online on 26 August and now appears in print (The French-Italian Public Consortium for Grapevine Genome Characterization *Nature* **449**, 463–467; 2007).

The grape variety concerned is Pinot Noir, the classic red grape of Burgundy. But the vine sequenced does not produce exactly the same grape as that grown in the vineyards. The consortium chose to sequence a variety called PN40024, which has been bred by successive self-crossings to reduce the high degree of sequence variation that is characteristic of all grapevine varieties. The inbred strain allows efficient assembly

of a high-quality sequence from whole-genome shotgun sequence data. In the shotgun technique the DNA is broken into many small fragments for sequencing and then reassembled from overlapping sequences.

The resulting genome sequence carries the imprint of millennia of selective breeding. For example, there are 116 genes and pseudogenes for terpene synthases, almost three times the number in the other three plant genomes so far sequenced. These enzymes synthesize the terpenoids that contribute to the aroma and flavour of wines, and pathways associated with tannins are similarly amplified.

Less obviously a target of selectivity are the genes that control the synthesis of resveratrol, the antioxidant credited with the health benefits claimed for moderate

consumption of red wine. Yet there is a modest expansion, compared with the other sequenced plants, of the stilbene synthase genes associated with resveratrol synthesis.

So can we look forward to genetically engineered 'designer' wines? Probably not. There is a market for new grapes, as exemplified by Cabernet Sauvignon clone 337, which is gaining ground in California's Napa Valley. But the flavour and aroma of wine depend on many other factors, such as growth conditions and production methods. And when it comes to producing wines with greater health-giving properties, the prospect sounds too good to be true. So it probably is. However, grapevines are notoriously susceptible to pathogens and stresses, such as drought, that other *Vitis* species can resist. The availability of this genome sequence should speed up progress on introducing the appropriate



WEGNER/NATURE PIC. LIB.

resistance into economically important varieties of *V. vinifera*.

With one grapevine genome sequenced, the way is clear for comparative oenogenomics. Yet when it comes to taste, perhaps the differences between a Pinot Noir with earthy and berry notes and a spicy or blowsy Gewürztraminer are best left to the realms of individual taste and a good thesaurus.

Charles Wenz

efficiently decouple. This is where the new work comes in^{1–3}, as it demonstrates how an on-chip photon field can act as a 'bus', or conduit, for quantum information, thus allowing a pair of distant qubits to interact at will.

Sillanpää and colleagues' qubits¹ are carefully designed, micrometre-sized elements of a superconducting circuit. These elements are coupled to opposite ends of an on-chip electrical resonator, or 'cavity', in which a standing electromagnetic wave of several millimetres' wavelength is established. Using an external source of microwave light, the authors first prepared one of their qubits in a superposition of its ground and first excited energy states. They then transferred this state to the photon field in the adjacent cavity — the bus. Finally, at the other end of the bus, they mapped the state of the photon field to the state of another qubit initially in its ground state. A further universal entangling operation between qubits could be achieved with extra tricks involving 'visits' to either higher qubit energy levels or larger numbers of photons. Such tricks are often applied in similar experiments in systems that use trapped-ion vibrations, rather than photons, as the bus medium^{4,5}.

Majer and colleagues' work² is similar, but comes with an additional twist. They also can carry out quantum-state transfer over a large distance, but in their case they never actually excite the intermediate photon field. Instead, they use 'virtual' photons, which are very weak perturbations of their cavity's quantum light field. This sleight-of-hand allows the authors to carry out a universal entangling operation

on a pair of distant qubits, without disturbing the bus itself.

Houck *et al.*³ performed complementary work by demonstrating a 'single-photon gun' that generates forwards-flying photons, instead of photons confined in a cavity, that have a well-defined phase of oscillation. They did this by first preparing an arbitrary quantum state in a superconducting qubit tightly coupled to a cavity. They then allowed the qubit to decay spontaneously, so that it emitted a single photon into a transmission line for microwave light. Convincing data from quantum-state tomography of both the qubit and the photon show how the qubit's initial state is transferred to the photon.

A great advantage of recent circuit architectures that exploit such 'cavity quantum electrodynamics' (cavity QED) approaches^{6–8} has been that the interaction between a superconducting qubit and a cavity can be much larger than the equivalent coupling between a real atom and a cavity. In addition, there is in principle room for hundreds of qubits on the same chip. Because any pair of qubits can be coupled, implementing algorithms and error-correction codes⁹ in a quantum computer will be significantly easier. One can also speculate that flying qubits such as those demonstrated by Houck *et al.*³ could be used to communicate between chips for a further scaling up.

But once qubits have been coupled using a photon field as a bus, how well can they be decoupled? One way to do this would be to ensure that the qubit excitation frequency is different from the resonance frequency of the

cavity. This can help to suppress the coupling between the qubits and the cavity, but might not be enough for long timescales and for many qubits. A possible future direction would be to combine the best features of cavity-bus architectures and 'nonlinear parametric couplers'¹⁰, which provide both tunable coupling and better isolation in their 'off' state.

A big remaining unknown is whether the lifetimes of complicated multi-qubit states can be made long enough for practical quantum computing. Nevertheless, the latest experiments^{1–3} on cavity-qubit interactions add significantly to the already large body of evidence showing that even relatively macroscopic objects can behave purely according to the laws of quantum physics — with all the promise that that holds for large-scale applications. ■

Antti O. Niskanen is at the VTT Technical Research Centre of Finland, POB 1000, 02044 VTT, Espoo, Finland. Yasunobu Nakamura is at the NEC Nano Electronics Research Laboratories, Tsukuba, Ibaraki 305-8501, and at the Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-0198, Japan.
e-mails: antti.niskanen@vtt.fi;
yasunobu@ce.jp.nec.com

1. Sillanpää, M. A., Park, J. I. & Simmonds, R. W. *Nature* **449**, 438–442 (2007).
2. Majer, J. *et al.* *Nature* **449**, 443–447 (2007).
3. Houck, A. A. *et al.* *Nature* **449**, 328–331 (2007).
4. Cirac, J. I. & Zoller, P. *Phys. Rev. Lett.* **74**, 4091–4094 (1995).
5. Schmidt-Kaler, F. *et al.* *Nature* **422**, 408–411 (2003).
6. Wallraff, A. *et al.* *Nature* **431**, 162–167 (2004).
7. Schuster, D. I. *et al.* *Nature* **445**, 515–518 (2007).
8. Astafiev, O. *et al.* *Nature* (in the press).
9. Shor, P. W. *Phys. Rev. A* **52**, R2493–R2496 (1995).
10. Niskanen, A. O. *et al.* *Science* **316**, 723–726 (2007).

Taking dendritic cells into medicine

Ralph M. Steinman¹ & Jacques Banchereau²

Dendritic cells (DCs) orchestrate a repertoire of immune responses that bring about resistance to infection and silencing or tolerance to self. In the settings of infection and cancer, microbes and tumours can exploit DCs to evade immunity, but DCs also can generate resistance, a capacity that is readily enhanced with DC-targeted vaccines. During allergy, autoimmunity and transplant rejection, DCs instigate unwanted responses that cause disease, but, again, DCs can be harnessed to silence these conditions with novel therapies. Here we present some medical implications of DC biology that account for illness and provide opportunities for prevention and therapy.

Immunology is a major force in medicine. It is needed to understand how prevalent diseases (Fig. 1) come about and how to develop preventions and treatments. This broad reach of the immune system reflects its two functions: to recognize diverse substances termed antigens, and to generate many qualitatively distinct responses. Dendritic cells (DCs), named for their probing, tree-like or dendritic shapes (from the greek 'dendron', meaning tree)¹ (Fig. 1), are pivotal for both recognition of a universe of antigens and control of an array of responses.

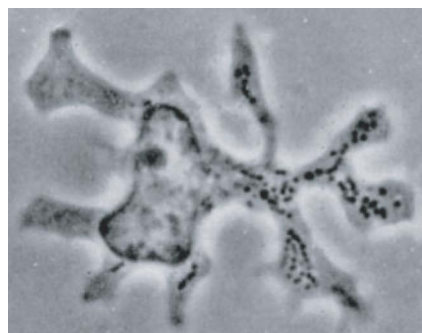
Previously, we reviewed some biological features of DCs². Here we illustrate medical implications of DCs, which control a spectrum of innate and adaptive responses. Innate immunity encompasses many rapid reactions to infection and other challenges³. Adaptive immunity, in contrast, is learned or acquired more slowly, in days to weeks; it has two hallmarks—exquisite specificity for antigens, and a durable memory to develop improved function on re-exposure to antigen. Adaptive responses are either immunogenic, providing resistance in infection and cancer, or tolerogenic, leading to silencing as is desirable in transplantation, autoimmunity and allergy. To date, the successes of immunology in the clinic have largely been based on antibodies made by B cells, but T-cell-mediated immunity, which has enormous, yet untapped, therapeutic potential, will be stressed here.

DCs are specialized to capture and process antigens *in vivo*^{4,5}, converting proteins to peptides that are presented on major histocompatibility complex (MHC) molecules and recognized by T cells. DCs also migrate to T-cell areas of lymphoid organs, where the two cell types interact to bring about clonal selection^{6–8}. The DC system is

thus designed to harness the recognition repertoire of T cells, consisting of billions of different lymphocytes, each with a distinct but randomly arranged antigen receptor. This repertoire in turn represents a virtually infinite 'drug library' for specific therapies that increase or decrease T-cell function.

Following clonal selection, DCs control many T-cell responses. Antigen-selected T cells undergo extensive expansion, a thousand-fold or more, as a result of division at a rate as high as 2–3 cell cycles a day⁹. Clones of lymphocytes are also subject to silencing or tolerance by so-called 'tolerogenic DCs'¹⁰, which either eliminate (delete)^{9,10} or block (suppress) T cells¹¹. If deletion is avoided, the clone undergoes differentiation to bring about an array of potential helper, killer and suppressive activities. For example, under the control of DCs, helper T cells acquire the capacity to produce powerful cytokines like interferon- γ to activate macrophages to resist infection by facultative and obligate intracellular microbes (T helper (T_H)1 cells)^{12–14}, or interleukin (IL)-4, -5 and -13 to mobilize white cells that resist helminths (T_H2 cells)¹⁵; or IL-17 to mobilize phagocytes at body surfaces to resist extracellular bacilli (T_H17 cells)¹⁶. Alternatively, DCs can guide T cells to become suppressive by making IL-10 (T regulatory (T_{reg})1 cells)¹⁷ or by differentiating into FOXP3⁺ cells¹¹. Finally, DCs induce the T-cell clone to acquire memory, allowing it to persist for prolonged periods and to respond rapidly to a repeated exposure to antigen^{18,19}.

There are four specialized features (Fig. 2) of DCs that contribute to their capacity to control T-cell recognition and responsiveness and, in turn, either prevent or generate disease. Recent advances in DC biology are outlined in Boxes 1–4. Briefly, (1) DCs are positioned to capture disease-causing antigens and to present these to lymphocytes in lymphoid organs^{20–22}, the sites for the generation of immunity and tolerance; (2) DCs have an endocytic system that is dedicated



- Infection
- Cancer
- Transplantation
- Autoimmunity and chronic inflammation
- Allergy
- Vaccines

Figure 1 | One of the first views of DCs in mouse spleen¹ (reproduced with permission from The Journal of Experimental Medicine; copyright 1973 Rockefeller University Press) and the conditions in which DCs and the immune system are most studied.

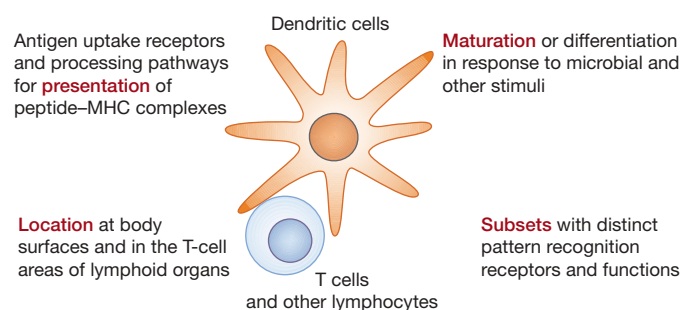


Figure 2 | Innate features of DCs, each an intricate subject.

¹The Rockefeller University, New York 10065, USA. ²Baylor Institute for Immunology Research, Dallas, Texas 75204, USA.

Box 1 | Location of DCs

DCs are a uniquely positioned, prime target for disease-relevant stimuli. DCs are abundant at body surfaces like the skin, pharynx, upper oesophagus, vagina, ectocervix and anus, and at so called internal or mucosal surfaces, such as the respiratory and gastrointestinal systems²¹. DCs actually extend their processes through the tight junctions of epithelia, which probably involves DC expression of the tight junction proteins claudins and occludins, without altering epithelial barrier function⁸⁶. This increases DC capture of antigens from the environment even when there is no overt infection or inflammation, probably allowing for the silencing of the immune system to harmless environmental antigens.

DCs at body surfaces can function locally, for example, to convert vitamins A and D to active retinoic acid and 1,25-dihydroxy-vitamin D₃. One consequence of the overlooked metabolic capacities of DCs is to increase the homing of immune cells to that mucosal surface and, in the case of retinoic acid, to help DCs differentiate suppressor T cells, which block autoimmune and inflammatory conditions⁸⁷.

After leaving peripheral tissues, DCs migrate with environmental, self and microbial antigens to lymphoid organs—a process that is guided by chemokines^{6,8} and can be enhanced by vaccination.

DCs have now been studied in intact lymphoid tissues without the need for cell isolation. These are the sites where immune resistance and tolerance are initiated. The DCs create a labyrinthine system within T-cell areas, while probing the environment through the continuous formation and retraction of processes²², and displaying antigens and other stimuli needed to initiate responses by appropriate clones of specific T cells^{88,89}.

New research reveals interactions of DCs in lymphoid tissues with other major classes of lymphocytes, B cells and natural killer cells⁵⁵.

Box 2 | Antigen presentation by DCs

DCs contain a specialized endocytic system^{4,5}, having many uptake receptors that deliver antigen to processing compartments^{5,24}. DCs then present peptides from processed proteins to CD4⁺ and CD8⁺ T cells, self and microbial glycolipids to NKT cells, and native antigens to B cells.

Many uptake receptors on DCs are lectins with carbohydrate-recognition capacity²⁸. Some, like LY75 (DEC-205/CD205) and mannose receptor MRC1/CD206, are type I transmembrane proteins with multiple contiguous lectin domains. A majority to date, such as DC-SIGN/CD209, langerin/CD207, ASGPR, OLR1/LOX-1 (ref. 40) and CLEC4A/DCIR (ref. 24), are type II proteins with a single external lectin domain. Among these, CD209 has attracted wide interest because it binds a number of microbes including HIV³⁵, dengue, cytomegalovirus, mycobacteria and candida, and because it can hinder DC maturation and contribute to immune evasion. Another lectin, langerin/CD207, is reported to degrade HIV in DCs, thereby reducing transmission to T cells.

Early literature on DCs took antigen uptake for granted, without realizing that ligation of uptake receptors increases the efficiency with which antigens are delivered to the immune system *in vivo* by ~100-fold^{9,19,38}. Targeting vaccines to these DC receptors should significantly improve the efficacy of T-cell mobilizing vaccines.

Following uptake of antigen, DCs are able to 'cross-present' antigens on MHC I to elicit CD8⁺ killer T cells^{33,90,91}. During cross-presentation, non-replicating protein antigens are internalized and somehow gain access to the cytoplasm before being processed by the proteasome for peptide presentation on MHC I. Critical steps seem to occur from less-acidic compartments. Cross-presentation allows DCs to induce CD8⁺ T cell responses to immune complexes, non-replicating forms of microbes and vaccines, and dying cells.

For dying cells, DC receptors for uptake remain to be defined *in situ*. Because cell death accompanies infection, cancer, transplantation, the normal turnover of self tissues and some viral vaccines, the uptake of dying cells is a starting point for DCs to capture antigens in many clinical settings.

Fc receptors, which recognize antigen-antibody complexes, mediate antigen uptake and both activating and inhibitory signals in DCs⁴⁷. For example, if inhibitory receptors are blocked, the binding of tumour cells coated with antibody leads to improved presentation of tumour antigens and production of the immune-stimulating cytokine IL-12 (ref. 47).

to antigen capture and processing, creating ligands for different classes of lymphocytes^{4,5}; (3) DCs differentiate or mature in response to a spectrum of stimuli²³ (Table 1), allowing them to bring about innate and adaptive responses that are potent and qualitatively matched to the disease-causing agent; and (4) DCs are comprised of subsets that differ from one another in terms of location, antigen presentation, and maturation^{24,25}. Location (Box 1) and antigen presentation (Box 2) allow DCs to efficiently select specific clones from the diverse recognition repertoire. Maturation (Box 3) and subsets (Box 4) allow DCs to control the diverse response repertoire of T cells and other classes of lymphocytes, such as B cells and natural killer cells.

Dendritic cells in infectious diseases

DCs induce resistance to infection. When microbial antigens are injected in association with DCs into mice, the animals acquire adaptive immunity to *Borrelia burgdorferi*, chlamydiae, *Leishmania major*, fungi, *Toxoplasma gondii*, malaria and HIV. Conversely, DC depletion reduces defences to viruses like CMV²⁶, HSV-2 and LCMV²⁷. In humans, a lack of circulating DCs during bacterial sepsis and dengue virus infection is associated with a poor prognosis.

A key concept is that DCs mature in distinct ways in response to different microbial components, thereby launching alternative versions of host immunity. The microbial ligands act on pattern recognition receptors³, including externally disposed Toll-like receptors (TLRs) and lectins²⁸, and the cytoplasmic NOD/NALP

family, RIG-I/DDX58 and MDA5/IFIH1 molecules. These pattern-recognition receptors can function synergistically^{14,29,30}.

In contrast, several microbes have the capacity to actively block DC maturation, for example, *Coxiella burnettii*, *Salmonella typhi*, anthrax lethal factor protein, *Plasmodia*, a *Mycobacterium ulcerans* mycolactone³¹ and viruses like vaccinia, herpes simplex, HIV, CMV, varicella zoster, HCV, Ebola/Marburg/Lassa fever and measles. An interesting exception is the effective attenuated yellow fever virus vaccine that may work by infecting and maturing DCs, allowing for antigen presentation to T cells³².

Furthermore, pathogens can alter other levels of DC physiology to evade an immune response. For example, the agent of plague, *Yersinia pestis*, and of typhoid fever, *Salmonella typhi*, selectively inject toxins into phagocytes, including DCs, and destroy the cells required for innate and adaptive protection. Influenza, measles and HSV-2 can induce apoptotic cell death in DCs³³. With some viruses, cell death occurs through the intermediate formation of giant cells or syncytia. CMV, herpes and *Mycobacterium tuberculosis* inhibit the migration step of DC function by blocking expression of CCR7 (ref. 34), a chemokine receptor that guides DCs into lymphatic vessels and onwards to lymphoid tissues^{6,8} (Box 1).

Microbes also can alter the function of DCs so that they switch T-cell responses from protective T_H1 to non-protective T_H2, as in infections with *Aspergillus fumigatus*, malaria and hepatitis C, or to IL-10 production in the case of *Bordetella pertussis*.

Table 1 | Some stimuli for dendritic-cell maturation

Stimulus	Details
Microbial products	Via Toll-like, NOD, RIG-I, Mda5 receptors
Lymphocytes (natural killer, NKT, $\gamma\delta$ T, $\alpha\beta$ T) and neutrophils	Via CD40, lymphotoxin $\alpha\beta$ receptors
Cytokines	TNF, interferons, TSLP and IL-10
Endogenous ligands	Uric acid, histamine, heat shock proteins, high mobility group box protein 1, defensins, ATP
Immune complexes	Via activating and inhibitory FcR

These stimuli induce distinct pathways of DC differentiation, whereupon the conditioned DCs activate distinct responses, for example, T_H1, T_H2, T_H17, T_{reg} cells. Two or more of these stimuli can act synergistically^{14,29,30}.

Box 3 | Maturation of DCs

DCs differentiate or mature in distinct ways in response to a spectrum of environmental and endogenous stimuli (Table 1). The maturation pathway then helps to designate which lymphocyte functions will be induced, and which products will be made by both DCs and lymphocytes^{23,92}.

In the steady state in the overt absence of maturation stimuli, DCs can induce tolerance when they capture self and environmental antigens⁹. Maintenance of tolerance can require PD-L1/CD274 (ref. 10) as well as FAS (ref. 93) on DCs. On infection or other causes of maturation, the DCs redeploy, but it is still not understood why DC differentiation is so rapid and extensive. One factor may be the high levels of required NF- κ B family proteins, with each family member(s) able to control different DC responses.

Maturing DCs can induce different types of CD4⁺ T cells (see text), such as T_H1, T_H2 or T_H17, to increase resistance. Other stimuli can yield 'tolerogenic' DCs, which induce Tr1 and FOXP3⁺ T_{reg} cells^{11,17,75} to silence immunity. Maturing DCs also express more IL-15 and activate inflammation and natural killer cells *in vivo*^{55,94}.

When DCs mature in response to microbial products, the expression of hundreds of genes is altered⁹⁵, leading to synthesis of cytokines, for example, IL-12 and type I interferons, which enhance innate and adaptive resistance^{26,96}. Although cytokines in turn can induce some components of DC maturation, the DCs that directly interact with microbial ligands are the immunologically more-active ones⁹⁷. The types of cytokines are influenced by the DC subset and the mode of DC activation. Several chemokines are also secreted in groups at a time⁹⁸, attracting different cells in succession to the site of DC maturation: phagocytes, memory lymphocytes and naive T cells⁹⁹. Numerous mechanisms also dampen the DC response to microbial products.

Maturation regulates antigen-processing by lowering the pH of endocytic vacuoles, activating proteolysis, and transporting peptide-MHC complexes to the cell surface⁵.

Importantly, maturing DCs remodel their surface, typically expressing many membrane-associated co-stimulatory molecules^{65,97}; these include members of the B7, TNF and Notch families.

A critical unknown is the definition of changes in DCs that link innate to adaptive immunity *in vivo*. The standard view that a combination of MHC-peptide ('signal one') and high B7-2/CD86 is sufficient to drive T-cell immunity is oversimplified. To influence T-cell differentiation, DCs additionally need to produce cytokines like IL-12 and type I interferons, or membrane-associated TNF family receptors—like CD40 (ref. 65) and lymphotoxin receptors—and TNF family members—like CD70 and OX40L to induce T_H1 (ref. 38) and T_H2 (ref. 100) cells, respectively.

At this time, there are no therapies that try to interrupt the microbial immune-evasion pathways that are summarized above.

Several microbes additionally can exploit DCs for purposes of replication and spread in the infected host. The lectin DC-SIGN/CD209 is used by dengue virus and Ebola virus to infect DCs. In the case of HIV-1, CMV and Ebola virus, the lectin additionally sequesters virus within DCs, which later transmit infectious virus to other targets, such as T cells³⁵. DCs are also implicated in the spread of varicella zoster, measles virus, poliovirus, *Aspergillus fumigatus*, LCMV, *Toxoplasma gondii*, prions and *Bacillus anthracis* spores.

All these consequences of the microbial-DC interaction have been analysed on myeloid DCs, mainly in tissue culture but not in patients. We expect that microbes also evade and exploit plasmacytoid DCs, which also help to resist pathogens³⁶ and are diminished in the blood during infections with HIV, HTLV-1 and RSV.

To counteract these mechanisms for pathogenesis of infectious disease, DCs are now being considered in the design of vaccines to prevent and treat infection by enhancing immunogenesis. A new concept is to deliver vaccine antigens to specific receptors on DCs (Box 2), along with stimuli to control DC maturation. For example, microbial proteins are genetically engineered into anti-receptor monoclonal antibodies, which then quickly and selectively target to

Box 4 | Subsets of DCs

There are several types of DCs, each with distinct markers and functions²⁵. Plasmacytoid DCs, so named because of cytologic similarities to antibody-producing plasma cells, can be involved in tolerance in their immature state. For maturation, these DCs selectively express activating Fc γ R as well as TLR7 and TLR9. When immune complexes containing DNA and RNA are bound and ingested, they signal potentially pathologic levels of type I interferons^{68,72}.

Other DCs, termed 'myeloid', also produce type I interferons. Myeloid DCs in different tissues can be further subdivided on the basis of expression of certain markers and functions. In the case of skin, DCs in the epidermis (Langerhans cells) express langerin/CD207 and DEC-205/CD205/LY75, and induce strong killer T-cell responses; some DCs in the dermis express DC-SIGN/CD209 and mannose receptor/CD206 and can activate antibody-forming B cells. Distinct skin DC populations also migrate to different areas of the draining lymph nodes¹⁰¹. Thus the outcome of skin vaccination may depend on the lectin and DC subset that pick up the vaccine.

In mouse spleen, a DC subset expresses DEC-205 and is particularly efficient for the cross-presentation of antigens on MHC I²⁴, including tumour cells, and also for the induction of IFN- γ -producing, T_H1 helper T cells^{12,13,38}. A second subset expresses DCIR2/CLEC4A4 and other lectins and is more efficient at processing antigens for presentation on MHC II²⁴. In lymph nodes draining mucosal tissues, the presence of an integrin, CD103/ITGAE, distinguishes a DC subset that cross-presents on MHC I¹⁰² but also synthesizes retinoic acid for the differentiation of FOXP3⁺ T_{reg} cells⁸⁷.

DC subsets communicate with each other, for example, plasmacytoid DCs produce interferons and membrane-bound co-stimulators that recruit other DCs to participate in immunity¹⁰³.

The production of most DC subsets is controlled in the steady state by the cytokine FLT3 ligand (FLT3LG)^{104,105}, whereas during inflammation and infection, another cytokine, GM-CSF, mobilizes increased numbers of monocyte-derived DCs¹⁰⁶. In the steady state, DCs in lymphoid tissues emanate from marrow progenitors in the blood¹⁰⁷ but not monocytes^{25,108}, although monocytes give rise to DCs in some non-lymphoid tissues in the steady state¹⁰⁸ and in many sites during inflammation²⁵.

large numbers of DCs within intact lymphoid tissues^{9,19,37}. The CD205 receptor, which is abundant on DCs in human lymphoid tissues, delivers antigen for processing onto both MHC class I and II (MHC I and II), increasing presentation efficiency >100-fold relative to non-targeted antigen^{19,38}. T_H1 responses, considered valuable for protection against many intracellular pathogens and tumours, also are induced when antigens from HIV, malaria, *Leishmania* and tumours are targeted to maturing CD205⁺ DCs^{19,37,38}. Significantly, these responses are broad; that is, they are capable of recognizing many peptides from a given microbial protein and in several MHC haplotypes^{19,39}, and the responses take place at mucosal surfaces—both key criteria in vaccine design. Other potential DC targets are being addressed, including LOX-1/OLR1 (ref. 40), MMR/CD206/MRC-1, DCIR/CLEC4A (ref. 24), DC-SIGN/CD209, toxin receptors and CD40.

Beyond the value of antigen-targeting, there is a need to mature DCs during vaccination in a way that is appropriate to the pathogen at hand. Here, one needs to define correlates of immunogenicity *in vivo*; that is, the specific changes in DCs required to generate protective responses from the response repertoire.

Dendritic cells in cancer

DCs are found in tumours in mice and patients. Yet tumours suppress immunity, especially locally and by many pathways⁴¹. Tumours express cytokines, like IL-6, vascular endothelial growth factor and IL-10, which suppress DCs through STAT3 signalling⁴². Tumours may condition local DCs to form suppressive T cells, such as FOXP3⁺ (ref. 43) and IL-13-producing CD4⁺ T cells (ref. 44) and natural killer T cells (NKT cells). DCs even support the clonogenic

growth of tumours in multiple myeloma. Therefore, like some infections, cancers have ways to evade and exploit DCs⁴⁵.

Nevertheless, immunology is providing treatments for cancer, mainly in the form of monoclonal antibodies. Antibodies, by virtue of their antigen-binding variable Fab regions, can block critical functions on cancer cells; and by virtue of their constant or Fc regions, antibodies can mobilize an attack by other Fc receptor (FcR)-bearing cells such as innate phagocytes and natural killer cells. Antibodies mediate the uptake and processing of tumour cells by DCs and also can trigger DC maturation (Table 1)^{46,47}. This portends the design of antibodies that harness select FcRs to induce better innate and adaptive anti-tumour immunity.

More emphasis is needed on cell-mediated immunity that also has a clear capacity to resist cancer⁴⁸. In phase I clinical research, adoptive transfer of killer T cells leads to regressions of melanoma and other cancers. There is also a major survival benefit from allogeneic bone marrow transplantation, in which lymphocytes from the marrow donor resist leukaemia and other haematologic malignancies (the 'graft versus leukaemia' reaction). In multiple myeloma, T cells that recognize glycolipids and peptides in the tumour are found in a premalignant stage of disease but are not found when the tumour grows out of control. In paraneoplastic diseases, T cells respond to antigens shared by the tumour and the nervous system, and probably resist the tumour but at the same time cause severe neurologic sequelae. Examination of colorectal cancers indicates that the presence of a T_H1-type immune response in the tumour correlates with a better prognosis. On the other hand, immune suppression predisposes to higher frequencies of several cancers. All of the above examples imply a role for immune surveillance by T cells against human cancer, as is also seen in mice⁴⁹.

DCs can be marshalled for the prevention and treatment of cancer for the following reasons: (1) tumours are replete with potential antigens, and they can become immunogenic when presented by DCs^{50,51}—this means that the immune attack on cancer can be broad enough to encompass multiple targets, including mutant proteins expressed by the cancer, and not just one target, where the latter favours immune escape; (2) likewise, DCs can activate and expand the different arms of cell-mediated resistance such as natural killer, NKT, $\gamma\delta$ T and $\alpha\beta$ T cells, each of which recognizes different alterations in cancer cells; and (3) DCs in systemic lymphoid organs and DCs generated *ex vivo* from progenitors in blood probably retain their immunizing capacities in cancer patients^{52,53}. Therefore one can test whether the specific features of DCs (Fig. 2) can be harnessed either to generate therapeutic immunity or to prevent cancer during premalignant or minimal residual disease stages.

Two DC-based immune therapy approaches are currently available. In one, DCs are generated *ex vivo*, loaded with tumour antigens, and re-injected to take advantage of the ability of DCs to migrate to the T-cell areas of lymphoid organs to induce strong T-cell and, perhaps, natural killer immunity^{54,55} (Fig. 3). Less explored is the evidence that DCs themselves can acquire killer activity for human tumours and express granzyme and perforin killer molecules⁵⁶. In another strategy, tumour cells or tumour antigens are targeted directly to DCs in the T-cell regions (Fig. 3), for example, within monoclonal anti-DC antibodies as discussed above.

Already a large number of phase I studies in humans have used the first approach, in which DCs are generated from precursors *ex vivo*. These early trials have only occasionally yielded significant tumour regressions^{57,58}, and there are no studies yet showing improved survival.

The current deficiencies in DC therapy need to be addressed in patients to deal with significant scientific and other obstacles. First, DC vaccines are being tried only in late-stage cancer patients who are immunosuppressed as a result of extensive radiation, chemotherapy and/or large tumour burdens. Why limit research to these patients, especially when DC therapy is nontoxic, including in the few patients who have experienced tumour regressions^{57,58}? Second, most injected

DCs remain at the injection site; only a few migrate to the draining lymphoid tissue, ~1% in mice and humans⁵⁹. Research in patients will be required to overcome this limitation. One example is to condition the injected site with cytokines like tumour necrosis factor (TNF), so that the injected DCs receive needed cues for migration into lymphatics. Third, more emphasis is needed on DC quality, by testing DCs for their capacity to induce helper and killer T cells with a high avidity for tumour antigens but few T_{reg} cells. Fourth, many initial vaccine studies used DCs charged with one or a few tumour antigens, whereas the potential of DCs rests with their still untapped capacity to elicit a strong and broad immune attack to lessen the chances of tumour escape. Fifth, vaccine studies need to be accompanied by in-depth immune monitoring to define assays for protective lymphocytes, the presence of which correlates with tumour regression and/or improved survival. For example, in HIV vaccines, it is thought that protective T cells will produce several and higher levels of cytokines like IFN- γ , IL-2 and TNF- α ; express low levels of the PD-1/CD279/PDCD1 regulatory protein; and have high functional avidity for antigen. Sixth, DC-based immune therapies require a coordinated research effort to generate DCs for clinical use and to help investigators systematically evaluate the many variables pertinent to efficacy.

Even with these limitations, early trials with *ex vivo* DCs have, in patients with advanced cancer, expanded T cells that recognize multiple tumour antigens⁶⁰ and that make the protective cytokine, IFN- γ (ref. 52). These observations provide a starting point to use DCs in a more concerted way to study their immune capacities in cancer patients, and to combine immunization with other therapeutic modalities. Chemotherapy⁶¹, radiotherapy and removal of tumour regulatory mechanisms, for example, CTLA4, B7-H4/VTCN1, IL-10, TGF- β (ref. 43), IL-13 (ref. 44), could be more specific and less toxic if combined with DC-based therapies.

The *ex vivo* technique is being developed so that it is easier to perform, including in collaborative studies. This approach allows the basic features of DCs (Fig. 2) to be controlled, for example: to load them appropriately with multiple antigens, including tumour cells^{57,58}, tumour RNA⁶², or, potentially, cancer stem cells; to select from many options for DC maturation (Box 3); and to specify the DC subset for clinical use⁶⁰ (Box 4).

A second DC-based strategy in cancer therapy would be to mobilize DCs directly within the patient, either within the tumour or within lymphoid organs. In mice, irradiated tumour cells injected intravenously are taken up by DCs. Alternatively, one can inject

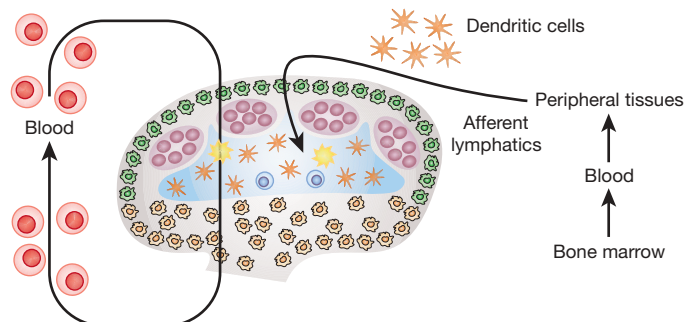


Figure 3 | DC localization in T-cell areas of immune organs such as lymph nodes, as shown here. Lymphocytes, including T cells, move from the blood into the node through special vessels (yellow). T cells (blue) remain in T-cell areas (blue), whereas B cells (pink) move to the B-cell areas (pink). DCs (orange stellate profiles) are primarily in the T-cell areas, whereas macrophages (green and orange profiles) are enriched in the periphery of the lymph node. To exploit this positioning in immunotherapy, DCs can be generated *ex vivo* and injected, whereupon homing via lymphatics brings small numbers of DCs to the T-cell areas (top). Alternatively antigens can be delivered directly to large numbers of DCs in the T-cell areas, for example, by incorporating the antigen within anti-DC monoclonal antibodies.

irradiated tumour cells that are transduced to express GM-CSF, which recruits the patient's DCs to capture and present tumour antigens^{63,64}. A new approach to the maturation of antigen-capturing DCs *in vivo* uses innate NKT cells that recognize glycolipids⁶⁵. The activated NKT cells mimic the effects of a combination of TLR and CD40 ligation on DCs and elicit long-lasting, protective CD4⁺ and CD8⁺ T-cell resistance. Another example is to selectively target cancer antigens within monoclonal antibodies to receptors on DCs, thus ensuring the delivery of antigen to large numbers of DCs in the T-cell areas of lymphoid organs, as opposed to the small numbers that successfully home from an injection site to these areas (Fig. 3).

Unfortunately, DC-mediated immunization against cancer is still an underdeveloped field. Yet this approach can exploit the patient's responses to cancer to produce more specific, less toxic, broader and longer-lasting therapies. There is an important unmet need to expand and coordinate research on DC-based therapies against proliferating and initiating cancer cells.

Dendritic cells in autoimmunity

Inappropriate responses to self constituents, in select genetic backgrounds, can lead to chronic inflammatory conditions, termed autoimmune diseases. DCs bearing self antigens are able to induce autoreactive T cells in mouse models of multiple sclerosis, cardiomyopathy and systemic lupus erythematosus.

It is increasingly appreciated that a pivotal step leading to human autoimmunity is an overproduction of a particular cytokine(s)⁶⁶, and subsets of DCs can be a major source. For example:

- TNF- α is a key cytokine in rheumatoid arthritis and other diseases, like psoriasis, and TNF- α blockade is a powerful therapy for many patients. In psoriasis, a major source of TNF- α is a DC subset infiltrating the affected skin⁶⁷.

- Lupus erythematosus, a systemic disease in which antibodies are formed against self constituents especially nucleoproteins, is accompanied by what is termed an 'interferon signature' in white blood cells⁶⁸. Type I interferon is made in large amounts by plasmacytoid DCs^{69,70}, which infiltrate the skin lesions of systemic lupus erythematosus. Viral nucleic acids, as well as self nucleoproteins internalized in the form of immune complexes, trigger TLR7 and TLR9, leading to type I interferon production. Moreover, the DNA-binding protein and cytokine, HMGB1, can deliver self DNA to TLR9 and ligate the RAGE/AGER receptor on DCs⁷¹. Interferon, in turn, drives the mobilization of activated granulocytes⁶⁸ and differentiates monocytes into DCs, which may present dying cells in an immunogenic rather than tolerogenic manner⁷². Ligation of activating FcR by the immune complexes that are formed during autoimmunity also drives DC maturation.

- Type I interferon and plasmacytoid DCs are also proposed to be pathogenic in other diseases, such as psoriasis⁷³, dermatomyositis and Sjögren's syndrome.

- IL-23 is a cytokine that drives disease, including psoriasis and inflammatory bowel disease⁷⁴, and DCs are major IL-23 producers.

Despite their role in inducing autoimmunity, DCs are also relevant to the therapy of these diseases: (1) various treatments, especially glucocorticosteroids, can reduce DC numbers and functions; (2) because DCs seem to be a source of pathogenic cytokines, the stimuli for cytokine production need to be identified and obviated; and (3) one potential antigen-specific strategy relates to the capacity of DCs to expand and induce T cells that suppress immunity, usually termed T_{reg} cells. T_{reg} cells can suppress other DCs that present disease-producing antigens. In the case of a spontaneous model of autoimmunity—diabetes in NOD mice—T_{reg} cells that recognize antigens in insulin-producing β cells can be generated by DCs and provide a therapeutic benefit even after the onset of disease^{11,75}.

Dendritic cells in allergy

DCs normally ensure the immune system is tolerant on exposure to harmless environmental antigens including the pollens, dust mites,

and foods that cause respiratory and intestinal allergy. The DCs at mucosal surfaces capture proteins and silence the corresponding T cells, possibly by expressing ICOSL, which in turn differentiates antigen-specific, IL-10-producing T_{reg} cells. In allergy, instead of this immune tolerance, CD4⁺ helper T cells develop along the atopic T_H2 pathway and yield interleukins (IL-4, -5 and -13) responsible for many aspects of disease, such as the formation of IgE-type antibodies. DCs both initiate the formation of pro-allergic T_H2 T cells and boost symptom-producing responses in already allergic mice⁷⁶.

How might the normal tolerizing roles of DCs be altered during allergy? Cytokines may again be pivotal, for example, recent studies reveal a role for thymic stromal lymphopoietin (TSLP) in atopic dermatitis⁷⁷ and asthma. TSLP matures human DCs from blood⁷⁷ and skin to elicit unusual T_H2 cells that secrete not only IL-4, -5 and 13 but also high levels of TNF. TSLP-treated mature DCs express OX40L/CD252/TNFSF4, a TNF family member. OX40L can instruct T cells to develop a T_H2-type of memory along with the production of allergic mediators like prostaglandin D₂.

Several new directions for allergy research involve DCs: (1) DC maturation by TSLP needs to be understood and blocked; likewise for TSLP production by epithelial cells, which is controlled by retinoid receptors as well as IL-1 and other inflammatory cytokines; (2) the substances that cause allergy may modify DC function directly by blocking IL-12 production by DCs and in turn favouring formation of allergic T_H2 cells⁷⁸; (3) in humans, new synthetic oligonucleotides can suppress the presentation of allergens by DCs to T_H2 cells; these compounds have shown promise in treating ragweed-induced allergic rhinitis; (4) DCs might be targeted to allow for formation of allergen-specific T_{reg} cells, which can treat allergy as observed in mouse models; and (5) new drugs can interfere with the pro-allergic functions of DCs and treat experimental asthma. The sphingosine-1-phosphate receptor agonist FTY720 blocks the migration of DCs from lung to lymph nodes where the DCs could present antigens to disease-causing T_H2 cells, whereas agonists for the D prostanoid 1 receptor condition DCs to induce disease-reducing T_{reg} cells⁷⁹.

Dendritic cells in transplantation

DCs play a key role in the outcome of organ and haematopoietic transplantation. DCs in grafted organs mature and migrate into the recipient, where they stimulate alloreactive T cells that bring about graft rejection. Recipient DCs also can capture portions of the graft^{80,81} and elicit organ rejection. Interestingly, suppressive drugs in current clinical use may act on the rejection-inducing DCs as well as the rejecting T cells⁸². In haematopoietic transplantation, recipient DCs are key initiators of T-cell-induced graft versus host reactions^{83,84}. Mechanisms have yet to be pinpointed to explain the maturation and migration of DCs that accompany transplantation.

It seems reasonable to propose that strategies to block these DC functions during transplantation will promote acceptance. However, recent discoveries show that DCs in grafted tissues can regenerate locally⁸⁵, thus providing a long-term source of antigen to stimulate rejection⁸⁴. In this light, alternative pathways to use DCs to induce transplantation tolerance are being assessed. One involves the activation of recipient natural killer cells, which reject donor DCs in tissue culture models and *in vivo*. Another pathway is to induce DCs to become tolerogenic, for example, to express the tolerogenic ILT3/LILRB4 molecule, or to induce graft-specific FOXP3⁺ T_{reg} cells to suppress graft-rejecting T cells⁸¹. In haematopoietic cell transplantation, strategies that lead to recipient DC depletion are currently being tested in the clinic.

Extending DC biology into medicine

Immunology, including T-cell-mediated immunity, has a central role in understanding how disease develops and in designing new treatments (Fig. 4). Here we suggest therapies aimed at the upstream events initiated by DCs. DCs can either intensify or subdue T-cell

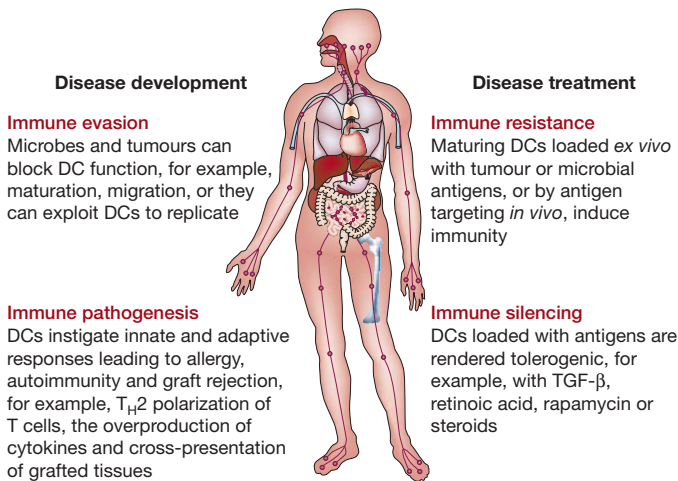


Figure 4 | Some potential roles for DCs in medicine.

responses, depending on whether resistance or tolerance needs to be increased.

Considerable evidence from studies in mice, which predominates in this review, shows how DCs act in a disease-specific manner. Increased emphasis on the antigens that elicit disease should provide medical approaches that are longer-lasting and less subject to side effects than current antigen non-specific ones. To repeat, DC-based analyses and therapies emphasize initial events in complex disease cascades, but more patient-based research is needed to take DCs into medicine.

The scientific rationale is that a patient's response to disease involves not only antigens and lymphocytes but also DCs that control antigen presentation and clonal selection (the immune recognition repertoire) as well as lymphocyte growth, differentiation and memory (the immune response repertoire). Each feature of DCs in Fig. 2 is an intricate area for research, the findings of which are facilitating a deeper analysis of disease and how to interrupt its progress. However, major gaps exist to take this science into medicine.

In mice, the current research emphasis has been on model antigens, isolated DCs, and highly selected populations of T cells with a single transgenic antigen receptor. Instead, research on DCs needs to be directed to the control of immune responses (1) *in situ* in intact animals, (2) within the natural immune repertoire, (3) to clinically relevant antigens, and (4) with new immune enhancers or adjuvants that act in defined ways on DCs. Even so, the immune systems of mice and men differ in many aspects. A demanding subject, but one which will enhance disease research, is the development of mice with immune systems derived from human sources.

More research needs to be done with patients. The patient sets the standards for the quality of knowledge that is required to understand many aspects of disease and its treatment. Often the pathogen (tumour, microbe, allergen, stimulus for autoimmune or autoinflammatory disease) is not easily or completely modelled in mice.

Although scientists and the public both desire 'translation' from mice to humans, there is an underappreciated need to create a sizeable limb of the scientific enterprise that will bring new methods, concepts and coordination to the basic study of disease with patients. This need is illustrated by the paucity of DC-based studies on the immunotherapy of cancer. Basic research with patients differs substantially from current outcome studies and drug-licensing trials, which test whether existing practices and concepts are clinically effective. In contrast, disease research yields new ideas and therapies. The biology of DCs is ready to be extended to dissect disease pathways and to direct its prevention and treatment. Despite the obstacles to research in patients, physicians and scientists have the knowledge and tools to think systemically about diseases, plan their therapies, and investigate how humans respond. DCs are an early

player in disease development and an unavoidable target in the design of treatments.

- Steinman, R. M. & Cohn, Z. A. Identification of a novel cell type in peripheral lymphoid organs of mice. I. Morphology, quantitation, tissue distribution. *J. Exp. Med.* **137**, 1142–1162 (1973).
- Banchereau, J. & Steinman, R. M. Dendritic cells and the control of immunity. *Nature* **392**, 245–252 (1998).
- Janeway, C. A. Jr & Medzhitov, R. Innate immune recognition. *Annu. Rev. Immunol.* **20**, 197–216 (2002).
- Guermonprez, P., Valladeau, J., Zitvogel, L., Thery, C. & Amigorena, S. Antigen presentation and T cell stimulation by dendritic cells. *Annu. Rev. Immunol.* **20**, 621–667 (2002).
- Trombetta, E. S. & Mellman, I. Cell biology of antigen processing *in vitro* and *in vivo*. *Annu. Rev. Immunol.* **23**, 975–1028 (2005).
- Cyster, J. G. Chemokines and the homing of dendritic cells to the T cell areas of lymphoid organs. *J. Exp. Med.* **189**, 447–450 (1999).
- Itano, A. A. & Jenkins, M. K. Antigen presentation to naive CD4 T cells in the lymph node. *Nature Immunol.* **4**, 733–739 (2003).
- Randolph, G. J., Angeli, V. & Swartz, M. A. Dendritic-cell trafficking to lymph nodes through lymphatic vessels. *Nature Rev. Immunol.* **5**, 617–628 (2005).
- Hawiger, D. *et al.* Dendritic cells induce peripheral T cell unresponsiveness under steady state conditions *in vivo*. *J. Exp. Med.* **194**, 769–780 (2001).
- Probst, H. C., McCoy, K., Okazaki, T., Honjo, T. & van den Broek, M. Resting dendritic cells induce peripheral CD8⁺ T cell tolerance through PD-1 and CTLA-4. *Nature Immunol.* **6**, 280–286 (2005).
- Luo, X. *et al.* Dendritic cells with TGF-β1 differentiate naive CD4⁺CD25[−] T cells into islet-protective Foxp3⁺ regulatory T cells. *Proc. Natl Acad. Sci. USA* **104**, 2821–2826 (2007).
- Pulendran, B. *et al.* Distinct dendritic cell subsets differentially regulate the class of immune responses *in vivo*. *Proc. Natl Acad. Sci. USA* **96**, 1036–1041 (1999).
- Maldonado-Lopez, R. *et al.* CD8α⁺ and CD8α[−] subclasses of dendritic cells direct the development of distinct T helper cells *in vivo*. *J. Exp. Med.* **189**, 587–592 (1999).
- Napolitani, G., Rinaldi, A., Berton, F., Sallusto, F. & Lanzavecchia, A. Selected Toll-like receptor agonist combinations synergistically trigger a T helper type 1-polarizing program in dendritic cells. *Nature Immunol.* **6**, 769–776 (2005).
- Seder, R. A., Paul, W. E., Davis, M. M., Fazekas de St. & Groth, B. The presence of interleukin 4 during *in vitro* priming determines the lymphokine-producing potential of CD4⁺ T cells from T cell receptor transgenic mice. *J. Exp. Med.* **176**, 1091–1098 (1992).
- Leibundgut-Landmann, S. *et al.* Syk- and CARD9-dependent coupling of innate immunity to the induction of T helper cells that produce interleukin 17. *Nature Immunol.* **8**, 630–638 (2007).
- Jonleitet, H., Schmitt, E., Schuler, G., Knop, J. & Enk, A. H. Induction of interleukin 10-producing, nonproliferating CD4⁺ T cells with regulatory properties by repetitive stimulation with allogeneic immature human dendritic cells. *J. Exp. Med.* **192**, 1213–1222 (2000).
- Badovinac, V. P., Messingham, K. A., Jabbari, A., Haring, J. S. & Harty, J. T. Accelerated CD8⁺ T-cell memory and prime-boost response after dendritic-cell vaccination. *Nature Med.* **11**, 748–756 (2005).
- Trumpfheller, C. *et al.* Intensified and protective CD4⁺ T cell immunity at a mucosal surface after a single dose of anti-dendritic cell HIV gag fusion antibody vaccine. *J. Exp. Med.* **203**, 607–617 (2006).
- Huang, F.-P. *et al.* A discrete subpopulation of dendritic cells transports apoptotic intestinal epithelial cells to T cell areas of mesenteric lymph nodes. *J. Exp. Med.* **191**, 435–442 (2000).
- Niess, J. H. *et al.* CX₃CR1-mediated dendritic cell access to the intestinal lumen and bacterial clearance. *Science* **307**, 254–258 (2005).
- Lindquist, R. L. *et al.* Visualizing dendritic cell networks *in vivo*. *Nature Immunol.* **5**, 1243–1250 (2004).
- Reis e Sousa, C. Dendritic cells in a mature age. *Nature Rev. Immunol.* **6**, 476–483 (2006).
- Dudziak, D. *et al.* Differential antigen processing by dendritic cell subsets *in vivo*. *Science* **315**, 107–111 (2007).
- Shortman, K. & Naik, S. H. Steady-state and inflammatory dendritic-cell development. *Nature Rev. Immunol.* **7**, 19–30 (2007).
- Dalod, M. *et al.* Interferon α/β and interleukin 12 responses to viral infections: pathways regulating dendritic cell cytokine expression *in vivo*. *J. Exp. Med.* **195**, 517–528 (2002).
- Probst, H. C. & van den Broek, M. Priming of CTLs by lymphocytic choriomeningitis virus depends on dendritic cells. *J. Immunol.* **174**, 3920–3924 (2005).
- Figdor, C. G., van Kooyk, Y. & Adema, G. J. C-type lectin receptors on dendritic cells and Langerhans cells. *Nature Rev. Immunol.* **2**, 77–84 (2002).
- Gautier, G. *et al.* A type I interferon autocrine-paracrine loop is involved in Toll-like receptor-induced interleukin-12p70 secretion by dendritic cells. *J. Exp. Med.* **201**, 1435–1446 (2005).
- Fritz, J. H. *et al.* Nod1-mediated innate immune recognition of peptidoglycan contributes to the onset of adaptive immunity. *Immunity* **26**, 445–459 (2007).
- Coutanceau, E. *et al.* Selective suppression of dendritic cell functions by *Mycobacterium ulcerans* toxin mycolactone. *J. Exp. Med.* **204**, 1395–1403 (2007).

32. Querec, T. *et al.* Yellow fever vaccine YF-17D activates multiple dendritic cell subsets via TLR2, 7, 8, and 9 to stimulate polyvalent immunity. *J. Exp. Med.* **203**, 413–424 (2006).
33. Albert, M. L., Sauter, B. & Bhardwaj, N. Dendritic cells acquire antigen from apoptotic cells and induce class I-restricted CTLs. *Nature* **392**, 86–89 (1998).
34. Khader, S. A. *et al.* Interleukin 12p40 is required for dendritic cell migration and T cell priming after Mycobacterium tuberculosis infection. *J. Exp. Med.* **203**, 1805–1815 (2006).
35. Geijtenbeek, T. B. H. *et al.* DC-SIGN, a dendritic cell specific HIV-1 binding protein that enhances trans-infection of T cells. *Cell* **100**, 587–597 (2000).
36. Colonna, M., Krug, A. & Cella, M. Interferon-producing cells: on the front line in immune responses against pathogens. *Curr. Opin. Immunol.* **14**, 373–379 (2002).
37. Boscardin, S. B. *et al.* Antigen targeting to dendritic cells elicits long-lived T cell help for antibody responses. *J. Exp. Med.* **203**, 599–606 (2006).
38. Soares, H. *et al.* A subset of dendritic cells induces CD4⁺ T cells to produce IFN- γ by an IL-12-independent but CD70-dependent mechanism *in vivo*. *J. Exp. Med.* **204**, 1095–1106 (2007).
39. Bozzacco, L. *et al.* DEC-205 receptor on dendritic cells mediates presentation of HIV gag protein to CD8⁺ T cells in a spectrum of human MHC I haplotypes. *Proc. Natl Acad. Sci. USA* **104**, 1289–1294 (2007).
40. Delneste, Y. *et al.* Involvement of LOX-1 in dendritic cell-mediated antigen cross-presentation. *Immunity* **17**, 353–362 (2002).
41. Gabrilovich, D. Mechanisms and functional significance of tumour-induced dendritic-cell defects. *Nature Rev. Immunol.* **4**, 941–952 (2004).
42. Kortylewski, M. *et al.* Inhibiting Stat3 signaling in the hematopoietic system elicits multicomponent antitumor immunity. *Nature Med.* **11**, 1314–1321 (2005).
43. Ghiringhelli, F. *et al.* Tumor cells convert immature myeloid dendritic cells into TGF- β -secreting cells inducing CD4⁺CD25⁺ regulatory T cell proliferation. *J. Exp. Med.* **202**, 919–929 (2005).
44. Aspod, C. *et al.* Breast cancer instructs dendritic cells to prime interleukin 13-secreting CD4⁺ T cells that facilitate tumor development. *J. Exp. Med.* **204**, 1037–1047 (2007).
45. Vicari, A. P., Caux, C. & Trinchieri, G. Tumour escape from immune surveillance through dendritic cell inactivation. *Semin. Cancer Biol.* **12**, 33–42 (2002).
46. Dhodapkar, K. M., Krasovsky, J., Williamson, B. & Dhodapkar, M. V. Anti-tumor monoclonal antibodies enhance cross-presentation of cellular antigens and the generation of myeloma-specific killer T cells by dendritic cells. *J. Exp. Med.* **195**, 125–133 (2002).
47. Kalergis, A. M. & Ravetch, J. V. Inducing tumor immunity through the selective engagement of activating Fc γ receptors on dendritic cells. *J. Exp. Med.* **195**, 1653–1659 (2002).
48. Blattman, J. N. & Greenberg, P. D. Cancer immunotherapy: a treatment for the masses. *Science* **305**, 200–205 (2004).
49. Shankaran, V. *et al.* IFN γ and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature* **410**, 1107–1111 (2001).
50. Figdor, C. G., De Vries, I. J., Lesterhuis, W. J. & Melief, C. J. Dendritic cell immunotherapy: mapping the way. *Nature Med.* **10**, 475–480 (2004).
51. Banchereau, J. & Palucka, A. K. Dendritic cells as therapeutic vaccines against cancer. *Nature Rev. Immunol.* **5**, 296–306 (2005).
52. Schuler-Thurner, B. *et al.* Rapid induction of tumor-specific type 1 T helper cells in metastatic melanoma patients by vaccination with mature, cryopreserved, peptide-loaded monocyte-derived dendritic cells. *J. Exp. Med.* **195**, 1279–1288 (2002).
53. Paczesny, S. *et al.* Expansion of melanoma-specific cytolytic CD8⁺ T cell precursors in patients with metastatic melanoma vaccinated with CD34⁺ progenitor-derived dendritic cells. *J. Exp. Med.* **199**, 1503–1511 (2004).
54. Fernandez, N. C. *et al.* Dendritic cells directly trigger NK cell functions: cross-talk relevant in innate anti-tumor immune responses *in vivo*. *Nature Med.* **5**, 405–411 (1999).
55. Lucas, M., Schachterle, W., Oberle, K., Aichele, P. & Diefenbach, A. Dendritic cells prime natural killer cells by trans-presenting interleukin 15. *Immunity* **26**, 503–517 (2007).
56. Stry, G. *et al.* Tumoricidal activity of TLR7/8-activated inflammatory dendritic cells. *J. Exp. Med.* **204**, 1441–1451 (2007).
57. Palucka, A. K. *et al.* Dendritic cells loaded with killed allogeneic melanoma cells can induce objective clinical responses and MART-1 specific CD8⁺ T-cell immunity. *J. Immunother.* **29**, 545–557 (2006).
58. O'Rourke, M. G. *et al.* Durable complete clinical responses in a phase I/II trial using an autologous melanoma cell/dendritic cell vaccine. *Cancer Immunol. Immunother.* **52**, 387–395 (2003).
59. De Vries, I. J. *et al.* Effective migration of antigen-pulsed dendritic cells to lymph nodes in melanoma patients is determined by their maturation state. *Cancer Res.* **63**, 12–17 (2003).
60. Banchereau, J. *et al.* Immune and clinical responses in patients with metastatic melanoma to CD34⁺ progenitor-derived dendritic cell vaccine. *Cancer Res.* **61**, 6451–6458 (2001).
61. Obeid, M. *et al.* Calreticulin exposure dictates the immunogenicity of cancer cell death. *Nature Med.* **13**, 54–61 (2007).
62. Gilboa, E. & Vieweg, J. Cancer immunotherapy with mRNA-transfected dendritic cells. *Immunol. Rev.* **199**, 251–263 (2004).
63. Soiffer, R. *et al.* Vaccination with irradiated autologous melanoma cells engineered to secrete human granulocyte-macrophage colony-stimulating factor generates potent antitumor immunity in patients with metastatic melanoma. *Proc. Natl Acad. Sci. USA* **95**, 13141–13146 (1998).
64. Laheru, D. & Jaffee, E. M. Immunotherapy for pancreatic cancer—science driving clinical progress. *Nature Rev. Cancer* **5**, 459–467 (2005).
65. Fujii, S., Liu, K., Smith, C., Bonito, A. J. & Steinman, R. M. The linkage of innate to adaptive immunity via maturing dendritic cells *in vivo* requires CD40 ligation in addition to antigen presentation and CD80/86 costimulation. *J. Exp. Med.* **199**, 1607–1618 (2004).
66. Banchereau, J. & Pascual, V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. *Immunity* **25**, 383–392 (2006).
67. Lowes, M. A. *et al.* Increase in TNF α and inducible nitric oxide synthase-expressing dendritic cells in psoriasis and reduction with efalizumab (anti-CD11a). *Proc. Natl Acad. Sci. USA* **102**, 19057–19062 (2005).
68. Bennett, L. *et al.* Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J. Exp. Med.* **197**, 711–723 (2003).
69. Siegal, F. P. *et al.* The nature of the principal type 1 interferon-producing cells in human blood. *Science* **284**, 1835–1837 (1999).
70. Asselin-Paturel, C. & Trinchieri, G. Production of type I interferons: plasmacytoid dendritic cells and beyond. *J. Exp. Med.* **202**, 461–465 (2005).
71. Tian, J. *et al.* Toll-like receptor 9-dependent activation by DNA-containing immune complexes is mediated by HMGB1 and RAGE. *Nature Immunol.* **8**, 487–496 (2007).
72. Blanco, P., Palucka, A. K., Gill, M., Pascual, V. & Banchereau, J. Induction of dendritic cell differentiation by IFN- α in systemic lupus erythematosus. *Science* **294**, 1540–1543 (2001).
73. Nestle, F. O. *et al.* Plasmacytoid dendritic cells initiate psoriasis through interferon- α production. *J. Exp. Med.* **202**, 135–143 (2005).
74. Hue, S. *et al.* Interleukin-23 drives innate and T cell-mediated intestinal inflammation. *J. Exp. Med.* **203**, 2473–2483 (2006).
75. Tarbell, K. V. *et al.* Dendritic cell-expanded, islet-specific, CD4⁺ CD25⁺ CD62L⁺ regulatory T cells restore normoglycemia in diabetic NOD mice. *J. Exp. Med.* **204**, 191–201 (2007).
76. Lambrecht, B. N. & Hammad, H. Taking our breath away: dendritic cells in the pathogenesis of asthma. *Nature Rev. Immunol.* **3**, 994–1003 (2003).
77. Soumelis, V. *et al.* Human epithelial cells trigger dendritic cell-mediated allergic inflammation by producing TSLP. *Nature Immunol.* **3**, 673–680 (2002).
78. Traidl-Hoffmann, C. *et al.* Pollen-associated phytoprostanes inhibit dendritic cell interleukin-12 production and augment T helper type 2 cell polarization. *J. Exp. Med.* **201**, 627–636 (2005).
79. Hammad, H. *et al.* Activation of the D prostanoid 1 receptor suppresses asthma by modulation of lung dendritic cell function and induction of regulatory T cells. *J. Exp. Med.* **204**, 357–367 (2007).
80. Inaba, K. *et al.* Efficient presentation of phagocytosed cellular fragments on the MHC class II products of dendritic cells. *J. Exp. Med.* **188**, 2163–2173 (1998).
81. Ochando, J. C. *et al.* Alloantigen-presenting plasmacytoid dendritic cells mediate tolerance to vascularized grafts. *Nature Immunol.* **7**, 652–662 (2006).
82. Hackstein, H. & Thomson, A. W. Dendritic cells: emerging pharmacological targets of immunosuppressive drugs. *Nature Rev. Immunol.* **4**, 24–34 (2004).
83. Shlomchik, W. D. *et al.* Prevention of graft versus host disease by inactivation of host antigen-presenting cells. *Science* **285**, 412–415 (1999).
84. Merad, M. *et al.* Depletion of host Langerhans cells before transplantation of donor alloreactive T cells prevents skin graft-versus-host disease. *Nature Med.* **10**, 510–517 (2004).
85. Merad, M. *et al.* Langerhans cells renew in the skin throughout life under steady-state conditions. *Nature Immunol.* **3**, 1135–1141 (2002).
86. Rescigno, M. *et al.* Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nature Immunol.* **2**, 361–367 (2001).
87. Coombes, J. L. *et al.* A functionally specialized population of mucosal CD103⁺ DCs induces Foxp3⁺ regulatory T cells via a TGF- β - and retinoic acid-dependent mechanism. *J. Exp. Med.* **204**, 1757–1764 (2007).
88. Bousso, P. & Robey, E. Dynamics of CD8⁺ T cell priming by dendritic cells in intact lymph nodes. *Nature Immunol.* **4**, 579–585 (2003).
89. Shakhar, G. *et al.* Stable T cell–dendritic cell interactions precede the development of both tolerance and immunity *in vivo*. *Nature Immunol.* **6**, 707–714 (2005).
90. Jung, S. *et al.* *In vivo* depletion of CD11c⁺ dendritic cells abrogates priming of CD8⁺ T cells by exogenous cell-associated antigens. *Immunity* **17**, 211–220 (2002).
91. Heath, W. R. *et al.* Cross-presentation, dendritic cell subsets, and the generation of immunity to cellular antigens. *Immunol. Rev.* **199**, 9–26 (2004).
92. Pulendran, B. Variations of the immune response with dendritic cells and pathogen recognition receptors. *J. Immunol.* **174**, 2457–2465 (2005).
93. Stranges, P. B. *et al.* Elimination of antigen-presenting cells and autoreactive T cells by Fas contributes to prevention of autoimmunity. *Immunity* **26**, 629–641 (2007).
94. Ohteki, T. *et al.* Essential roles of DC-derived IL-15 as a mediator of inflammatory responses *in vivo*. *J. Exp. Med.* **203**, 2329–2338 (2006).
95. Granucci, F. *et al.* Inducible IL-2 production by dendritic cells revealed by global gene expression analysis. *Nature Immunol.* **2**, 882–888 (2001).
96. Reis e Sousa, C. *et al.* *In vivo* microbial stimulation induces rapid CD40L-independent production of IL-12 by dendritic cells and their re-distribution to T cell areas. *J. Exp. Med.* **186**, 1819–1829 (1997).

97. Sporri, R. & Reis e Sousa, C. Inflammatory mediators are insufficient for full dendritic cell activation and promote expansion of CD4⁺ T cell populations lacking helper function. *Nature Immunol.* **6**, 163–170 (2005).
98. Sallusto, F. *et al.* Rapid and coordinated switch in chemokine receptor expression during dendritic cell maturation. *Eur. J. Immunol.* **28**, 2760–2769 (1998).
99. Piqueras, B., Connolly, J., Freitas, H., Palucka, A. K. & Banchereau, J. Upon viral exposure, myeloid and plasmacytoid dendritic cells produce 3 waves of distinct chemokines to recruit immune effectors. *Blood* **107**, 2613–2618 (2006).
100. Ito, T. *et al.* TSLP-activated dendritic cells induce an inflammatory T helper type 2 cell response through OX40 ligand. *J. Exp. Med.* **202**, 1213–1223 (2005).
101. Kissenpfennig, A. *et al.* Dynamics and function of Langerhans cells *in vivo*: dermal dendritic cells colonize lymph node areas distinct from slower migrating Langerhans cells. *Immunity* **22**, 643–654 (2005).
102. Del Rio, M. L., Rodriguez-Barbosa, J. I., Kremmer, E. & Forster, R. CD103⁺ and CD103⁺ bronchial lymph node dendritic cells are specialized in presenting and cross-presenting innocuous antigen to CD4⁺ and CD8⁺ T Cells. *J. Immunol.* **178**, 6861–6866 (2007).
103. Yoneyama, H. *et al.* Plasmacytoid DCs help lymph node DCs to induce anti-HSV CTLs. *J. Exp. Med.* **202**, 425–435 (2005).
104. D'Amico, A. & Wu, L. The early progenitors of mouse dendritic cells and plasmacytoid predendritic cells are within the bone marrow hemopoietic precursors expressing Flt3. *J. Exp. Med.* **198**, 293–303 (2003).
105. Karsunky, H., Merad, M., Cozzio, A., Weissman, I. L. & Manz, M. G. Flt3 ligand regulates dendritic cell development from Flt3⁺ lymphoid and myeloid-committed progenitors to Flt3⁺ dendritic cells *in vivo*. *J. Exp. Med.* **198**, 305–313 (2003).
106. Leon, B., Lopez-Bravo, M. & Ardavin, C. Monocyte-derived dendritic cells formed at the infection site control the induction of protective T helper 1 responses against *Leishmania*. *Immunity* **26**, 519–531 (2007).
107. Liu, K. *et al.* Origin of dendritic cells in peripheral lymphoid organs of mice. *Nature Immunol.* **8**, 578–583 (2007).
108. Varol, C. *et al.* Monocytes give rise to mucosal, but not splenic, conventional dendritic cells. *J. Exp. Med.* **204**, 171–180 (2007).

Acknowledgements The authors thank C. Moberg and M. Nussenzweig for extensive comments and J. Adams for assistance with the manuscript. We are grateful to our patients and colleagues, for their many contributions, and to the NIH and several foundations for support. A version of the review with a more complete bibliography is published on the authors' websites: www.biiir.org and www.rockefeller.edu/labheads/steinman/generalReviews.php.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence should be addressed to R.M.S. (steinma@mail.rockefeller.edu).

Predicting evolutionary patterns of mammalian teeth from development

Kathryn D. Kavanagh^{1†}, Alistair R. Evans¹ & Jukka Jernvall^{1,2}

One motivation in the study of development is the discovery of mechanisms that may guide evolutionary change. Here we report how development governs relative size and number of cheek teeth, or molars, in the mouse. We constructed an inhibitory cascade model by experimentally uncovering the activator–inhibitor logic of sequential tooth development. The inhibitory cascade acts as a ratchet that determines molar size differences along the jaw, one effect being that the second molar always makes up one-third of total molar area. By using a macroevolutionary test, we demonstrate the success of the model in predicting dentition patterns found among murine rodent species with various diets, thereby providing an example of ecologically driven evolution along a developmentally favoured trajectory. In general, our work demonstrates how to construct and test developmental rules with evolutionary predictability in natural systems.

A recurring promise of evolutionary developmental biology is the discovery of the mechanisms and rules that govern the production of the phenotypic variation available for natural selection^{1–7}. In many cases, independent evolutionary acquisitions of morphological similarities have been linked to parallel changes in the genome^{8–11}, but selection may also produce the same phenotypic result in different evolutionary trials by altering different genes or interactions^{12–15}. Whereas in the former cases gene-level or low-level rules may have predictive power across a broad range of taxa, in the latter cases models may have to be built on a higher level of organization. Regardless of the organizational levels invoked, explicit or inferred developmental rules with evolutionary relevance must be shown to favour specific evolutionary trajectories, and previous demonstrations have ranged from theoretical and computational models to selection and developmental experiments^{1–7}. Advances in developmental genetics now allow us to identify mechanisms that may bias the production of phenotypic variation, which in part will help to move evolutionary developmental biology into the realm of a predictive science.

The mammalian dentition is a classic system in which developmental mechanisms have been used to explain variation in shape and size^{16–22}. The timing of molar initiation during development has been extensively studied in primates, not least because molar proportions are used as a diagnostic feature in palaeontology. Differences in the timing of molar initiation, mineralization and eruption as well as in molar size and number have been linked to species-specific traits in diet, life history, maturation and brain size^{23–26}. In addition, regulation of molar size and number continues to have medical relevance in connection with human third molars, or wisdom teeth, which are frequently surgically removed with a risk of complications²⁷.

Mammalian molars develop sequentially in an anterior to posterior direction (Fig. 1a), resembling the development of segmental structures, but it remains unknown how molar initiation or size is regulated along the tooth row. Mechanisms including available space in the jaw and inhibition between developing teeth have both been proposed to regulate molar initiation^{21,22,25}. Because experimental evidence and mathematical modelling have implicated a balance of molecular signals activating and inhibiting the formation of

teeth^{28–30}, here we examine whether inhibitory dynamics could explain the initiation and size of adjacent molars in mouse (*Mus musculus*) and whether these dynamics can account for aspects of evolutionary patterns of teeth.

Inhibitory dynamics of molar initiation

As in most eutherian mammals, mice have three molars that develop sequentially over several days³¹. The development of each individual tooth is punctuated by the formation of the epithelial signalling

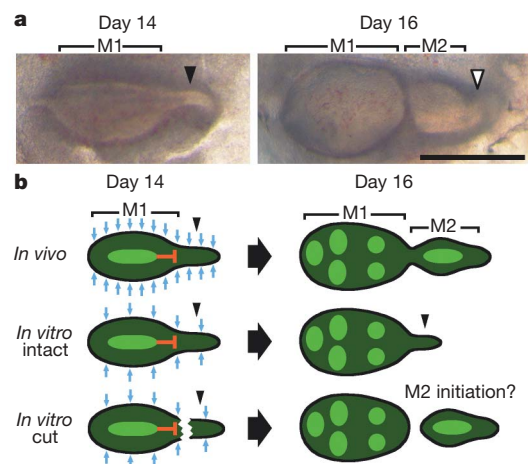


Figure 1 | Hypotheses on the sequential initiation and inhibition of mammalian cheek teeth. **a**, Mouse molars develop sequentially, and the dental lamina extending posteriorly (black arrowhead) from the developing M1 gives rise to M2 at day 16. M3 forms (white arrowhead) posterior to M2 about ten days later. **b**, In comparison with the situation *in vivo*, M1 development proceeds normally *in vitro* and the secondary enamel knots form at day 16 (bright green). In contrast, M2 initiation is delayed *in vitro*. We suggest that this delay is due to a decrease in mesenchymally secreted activators (blue arrows), whereas M1 continues to inhibit M2 normally. To test this, we cut the posterior tail that forms M2 from M1. Anterior is towards the left. Scale bar, 0.5 mm.

¹Evolution & Development Unit, Institute of Biotechnology, PO Box 56 (Viikinkaari 9), FIN-00014 University of Helsinki, Finland. ²Department of Ecology and Evolution, Stony Brook University, Stony Brook, New York 11794, USA. †Present address: School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, New York 11794, USA.

centres, the enamel knots²⁸. A primary enamel knot forms at the onset of the tooth crown development, followed by secondary enamel knots that appear at the future positions of major molar features, the cusps. Because mutations affecting the inhibition of enamel knots can have fused or extra cusps and molars³⁰, we postulate that the first developing molar could inhibit the development of subsequent molars, an effect that we also propose to be accentuated by culture conditions. In culture, although the first molar (M1) develops at essentially the same rate as that *in vivo*, posterior molars are frequently delayed by several days or never develop at all. The culture conditions, which involve the dissection of tooth germs from surrounding tissue, seem to disrupt the mesenchymal influence on the balance of activator and inhibitor molecules regulating molar development (Fig. 1b).

To test these ideas, we cultured lower first molar tooth germs from a mouse that expressed green fluorescent protein (GFP) fused with Cre-recombinase in the Sonic Hedgehog (*Shh*) locus (hereafter called *ShhGFP* mice). In the developing tooth crown, *Shh* is first upregulated only in the cells of the enamel knots; later, during differentiation, *Shh* expression is detected in the enamel-secreting ameloblasts throughout the crown^{28,32}. Because enamel knots are difficult to detect under normal culture conditions, the epifluorescence of *ShhGFP* mice allowed us to pinpoint the future positions of the molars and cusps *in vitro*, thereby permitting us to follow the sequential odontogenesis continuously (Fig. 2a). The *ShhGFP* construct is a *Shh*-null allele; we therefore cultured heterozygous *ShhGFP* molars. Similarly to a previous report on limb development³³, we found the tooth development and morphology of heterozygous *ShhGFP* mice to be normal. We also examined the development of wild-type molars, and the pattern of results remained essentially the same.

Using a standard Trowell culture system³⁴, we first cultured *ShhGFP* molars starting from embryonic day 14, at which time the M1 primary enamel knot has formed. We cultured both intact tooth germs and teeth in which we surgically separated the developing M1 from its posterior tail that is fated to give rise to the second (M2) and third (M3) molars (Figs 1 and 2a). Cultures were monitored daily, and the initiation of each tooth was reconstructed from time-lapse images (Fig. 2a). For the cultured intact tooth germs, the results show that only 11% of explants formed M2 enamel knots after two days in culture, a period equivalent to the timing of M2 initiation *in vivo* (Fig. 2b). Additional intact

explants developed M2s during the subsequent days, and 54% of M2s were initiated by 12 days in culture (Fig. 2b). In contrast to this delayed initiation of M2 development, M1 development progressed at a fairly normal rate and all M1s had formed secondary enamel knots by three days into the culture (Fig. 2a), matching the rate of development *in vivo*²⁸. The normal development of M1 implies that nutritional deficiency is unlikely to cause the delay in the posterior molars, but it supports the hypothesis of inhibition by M1 (Fig. 1b).

For the explants in which the tail had been cut off from the rest of the tooth germ, the results show that 98% of the separated tails formed M2s, with 68% of them occurring at the *in vivo* rate (Fig. 2b). Therefore, rather than inflicting irreversible damage on the small posterior bud, the separation seems to rescue M2 development from an inhibitory effect of M1. We interpret this result to also exclude an inhibitory gradient going through the jaw and teeth, increasing from anterior to distal, because in that case we would not expect the separation to rescue M2s. Furthermore, in almost half of the cut explants, M3 development was initiated, often before expected M3 initiation *in vivo* (Fig. 2c).

At the day 14 cap stage, when the M1 enamel knot has formed, M1 expresses the genes encoding several signalling molecules²⁸, including diffusible inhibitors. Of these, at least *ectodin* (also known as *Sostdc1* and *wise*, inhibitor of bone morphogenetic proteins (BMPs) and Wnts), *Bmp3* and *folliculin* (both encoding inhibitors of Activin A and BMPs) are strongly expressed in the enamel knot or anterior portion of the day 14 M1 (refs 30, 35, 36). Therefore, to test how an earlier release from inhibition affects posterior molars, we cut the posterior tails also at day 13, when the M1 primary enamel knot would only just be forming. These results show that posterior molar initiation was accelerated further: 90% of M2s were now initiated one day earlier than *in vivo* (Fig. 2b). The initiation of M3 development was also markedly accelerated (Fig. 2c). In addition, in one of the explants, a fourth molar (M4) formed seven days into the culture. We note that even though the tails giving rise to M3s were too small to be dissected from M2s, molar initiation was always sequential and in no case did we observe a simultaneous initiation of M2 and M3. Thus, M3 initiation is likely to be inhibited by M2 and, consequently, M4 is inhibited by M3.

Our results indicate that, as seems to occur with the regulation of fibroblast growth factors during tooth development³⁷, the balance

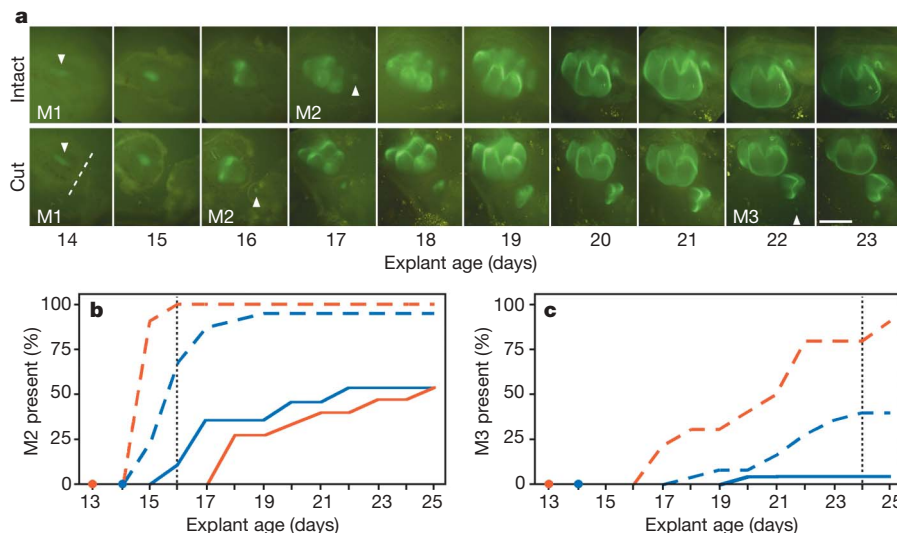


Figure 2 | Posterior molars are initiated earlier *in vitro* when separated from M1. **a**, The epifluorescence of cultured *ShhGFP* teeth allows daily monitoring of the enamel knots to test whether cutting the posterior tail (dashed line) accelerates molar initiation (white arrowheads). **b**, **c**, Cumulative percentage curves show that, in comparison with the intact explants (solid lines), the cut explants (dashed lines) at day 14 (blue) and day

13 (red) have an accelerated initiation of M2 (**b**) and M3 (**c**). *In vivo* M2 and M3 initiation times are marked with dotted vertical lines. Mann–Whitney *U*-tests on M2 and M3 age differences between intact and cut explants after 12 days of culture are all $P \leq 0.001$ (see Supplementary Information). $n = 28$ and $n = 25$ for day 14 intact and cut explants, respectively, and $n = 15$ and $n = 10$ for day 13 intact and cut explants, respectively. Scale bar, 0.5 mm.

between enamel knot activation and inhibition may be more important for tooth initiation than the absolute magnitude of signals themselves. Initially, the *in vitro* culture seems to decrease the level of mesenchymal activators required for M2 induction whereas removal of the inhibitory effect of M1 restores the inductive balance (Figs 1, 2). One obvious assumption linked to these interpretations is that culture conditions decrease mesenchymal activators required for enamel knot formation (Fig. 1b). To test this, we explored the effects of BMP4 and Activin A by using protein-releasing bead experiments. Both *Bmp4* and *activin βA* are intensely expressed in the mesenchyme at the onset of primary enamel knot formation, and both have been implicated as mediators of epithelial–mesenchymal induction events leading to the formation of enamel knots^{30,36,38–40}. We placed beads releasing BMP4 or Activin A immediately distally to intact day 14 tooth germs. The results show that both molecules are individually able to accelerate the formation of M2s, although not to the extent that separation from M1 achieved (Fig. 3).

Taken together, our experimental results suggest that the initiation timing of posterior molars depends on previous molars through a dynamic balance between intermolar inhibition and mesenchymal activation. Because of the importance of molar size in evolution^{6,16–29}, we next explored how these developmental dynamics might bias the production of phenotypic variation available for natural selection.

Molar initiation and size

To link our results for the process of molar initiation to morphological patterns, we measured from our experiments how tooth size is affected by changes in tooth initiation (see Methods and Supplementary Information). The results show that the removal of inhibition on posterior molars results not only in earlier tooth initiation but also in larger posterior teeth. After 12 days of culture, M2s in the cut explants were twice the size of M2s in the intact explants (cut versus intact day 14 means are 0.27 and 0.13 mm², $P < 0.001$, and day 13 means are 0.23 and 0.11 mm², $P < 0.001$; Mann–Whitney *U*-tests). Furthermore, the cut explant M2s are larger not only as a result of earlier

initiation but also because they grow faster (see Supplementary Information). In contrast, M1 sizes have marginally decreased in the cut explants, suggesting that dissection caused disruption and also that inhibition is always from anterior to posterior (cut versus intact day 14 means are 0.57 and 0.67 mm², $P = 0.027$, and day 13 means are 0.39 and 0.46 mm², $P = 0.123$; Mann–Whitney *U*-tests). Nevertheless, in comparison with the intact explants, both the day 13 and day 14 cut explants produced 15–38% more ‘tooth’, measured as the sum of the molar surface areas ($P = 0.005–0.012$; Mann–Whitney *U*-tests).

In the intact day 13 explants, the initiation of posterior molars was delayed in comparison with that of intact day 14 explants (Fig. 2b, c) and the M1s were also smaller, perhaps because of decreased mesenchymal activation that limited development at this earlier stage. Despite this typical retardation of day 13 tooth development *in vitro* (J. Jernvall and K. D. Kavanagh, unpublished observations), sizes of the day 13 cut explant M2s matched, and that of the M3s exceeded, the sizes of corresponding teeth from the day 14 cut explants (Fig. 4a; see Supplementary Information). The earlier separation from M1 therefore seems to lead to a tendency in which molar sizes become more equal (Fig. 4a). Whereas, for example, the day 14 M3s could in principle catch up with the day 13 M3s, this would require the former to grow more than twice as long as the latter. We consider this situation unlikely because in our cultures the onset of mineralization seemed to be the same in both the day 13 and day 14 cut explants.

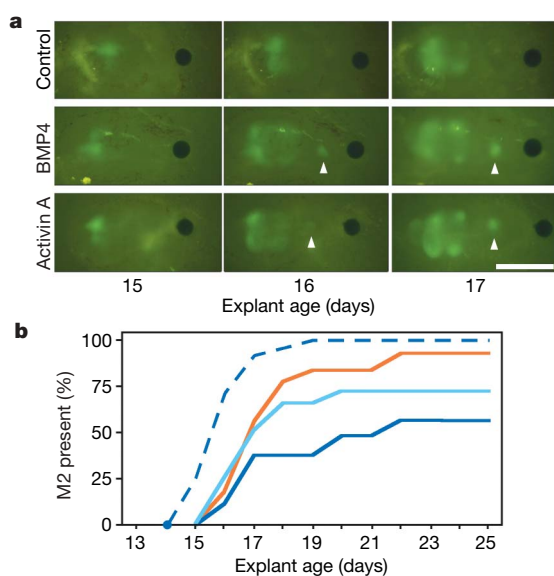


Figure 3 | Initiation of posterior molars can be stimulated by mesenchymal activators. **a**, Protein-releasing beads were placed posteriorly to day 14 explants and the initiation of M2s was monitored (white arrowheads). **b**, Both BMP4- (light blue line) and Activin A- (orange line) releasing beads accelerate M2 initiation, falling between the intact (solid blue line) and cut explants (dashed blue line). Mann–Whitney *U*-tests on M2 age differences between protein and control explants after 12 days of culture are $P = 0.122$ ($n = 16$) for BMP4 and $P = 0.014$ ($n = 19$) for Activin A explants (see Supplementary Information). Scale bar, 0.5 mm.

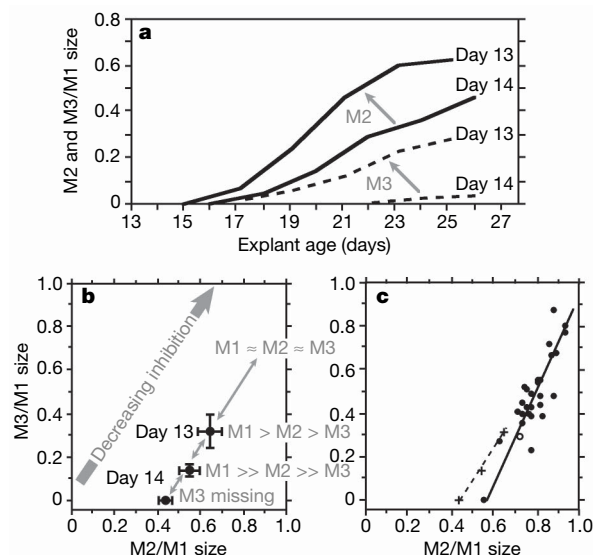


Figure 4 | From molar initiation to predicting molar proportions in murine species. **a**, Removal of inhibition results in earlier initiation and more equal-sized posterior molars. **b**, Changes in inhibition provide a trajectory through the morphospace in which more equal-sized molars are found with low inhibition (day 13 explants; error bars denote s.e.m.). In contrast, increasing inhibition (day 14 explants) leads to smaller posterior molars and eventually the lack of M3. **c**, The molar proportions of 29 species of murine rodents (black circles; *Mus musculus* is marked with an open circle) fall close to the experimental data (crosses and dashed line). We note the lack of M3 when M2 is about half the size of M1, in both the experimental and the macroevolutionary data. For the experimental data, the slope drawn through the means of day 14 and day 13 molar sizes is 1.848 and the intercept is -0.833 . When the 12 cut explants without M3s (all except one were day 14) are plotted separately (**b**), the resulting reduced major axis regression slope is 1.519 and the intercept is -0.673 . When M1 sizes just before they reach their asymptotic sizes are used, approximating the growth stage of measured M2s and M3s, the reduced major-axis regression slope is 2.024 and the intercept is -0.997 . For the macroevolutionary data (**c**), the reduced major-axis regression slope is 2.150 and the intercept is -1.219 ($r^2 = 0.740$). For details see Supplementary Information.

An inhibitory cascade model

The inhibitory dynamics (Figs 2, 3) and shifting molar proportions (Fig. 4a) are indicative of an inhibitory cascade, or a 'ratchet' in which subsequently developing teeth are cumulatively affected by previous developmental events. The inhibitory cascade can be formalized as a simple high-level model in which a balance between activation and inhibition results in equal-sized molars ($M1 \approx M2 \approx M3$) and increasing inhibition has a cumulative effect on the posterior teeth giving a distinct $M1 > M2 > M3$ pattern (Fig. 4b). The relative molar sizes determined by the model can be stated as $1 + [(a - i)/i](x - 1)$, in which, at each molar position (x), tooth size results from the relative strengths of activators (a) and inhibitors (i). As a result of the ratcheting nature of the inhibition, a change in inhibition (or activation) affects the relative size of M3 more than that of M2 (Fig. 4b). Nevertheless, molars have shared covariance patterns, so the relative size of adjacent teeth allows one to predict the presence and size of additional teeth. For example, M3s are missing when M2 size falls below half that of M1 (Fig. 4b). Conversely, our case of M4 occurred when the size of M3 equalled that of M2 (Supplementary Information), perhaps indicating that the evolution of supernumerary teeth is most likely when tooth activation and inhibition are in balance.

A macroevolutionary test of the model

Because our model makes broad predictions about the relative sizes of individual teeth, to test the model we focused on a sample of 29 species of murine rodents covering a wide spectrum of ecological adaptations and phylogenetic lineages representative of the entire subfamily^{41,42} (Supplementary Information). Tooth rows were digitized with a high-resolution laser scanner and the molar crown areas were measured with the MorphoBrowser database containing the three-dimensional tooth scans⁴².

The basic prediction from the experiments is that with an increase in relative size of M2, M3 should increase more. The results show that molar proportions follow this expectation closely (Fig. 4c), although the macroevolutionary patterns seem to show a slightly greater increase in posterior molars than the experimental prediction (Fig. 4c). We suspect that this is because our developmental data were derived from cultured teeth *in vitro* in which M1 was near mineralization whereas M2 and especially M3 could grow further, increasing their relative sizes. Indeed, when ante-asymptotic M1 sizes are used for the experimental data, the slopes of the molar size relationships are very similar between the experiments (2.02) and species (2.15). Conversely, in our molar diversity data, we have one species, golden-bellied water rat (*Hydromys chrysogaster*), which lacks M3 altogether. Matching the prediction from mouse explants lacking M3s, M2 in *Hydromys* is about half the size of M1 (Fig. 4c). Thus, despite the limitations of *in vitro* cultures (uncut M3s and incomplete differentiation), these results may implicate the inhibitory cascade in regulating tooth proportions.

Next, to test how closely the macroevolutionary data follow the explicit prediction of the inhibitory cascade model $1 + [(a - i)/i](x - 1)$, we first calculated the predicted sizes of M3s on the basis of the relative size of M2s (see Fig. 5a and Methods). Both the slope (2.0) and the intercept (-1.0) of the model prediction are within the 95% confidence intervals of the macroevolutionary data. To examine further the consistency of the tooth-to-tooth inhibitory relay in our data, we generated a random relay model, in which the strength of inhibition changed between teeth, by randomly reshuffling the M2-based predictions of M3 sizes 1,000 times (see Fig. 5a, Methods and Supplementary Information). The results show that whereas the random relay still produces correlated variation between relative M2 and M3 sizes (because, for example, it is unlikely that a large M2 is followed by a very small M3), its predictions are not congruent with our macroevolutionary data or model (Fig. 5a and Supplementary Information). We interpret these results as further implicating the

inhibitory cascade as a 'ratchet' generating predictable size differences along the molar row.

One phenotypic outcome of the ratchet is the high variability of M3, a result that agrees well with data from populations and species^{18–21,43,44}. Whereas the high variability of M3 has been linked to available space in the jaw and difficulty in measuring small M3s, the inhibitory cascade may provide null expectations for M3 variability. Another phenotypic result specific to the model is that M2 makes up roughly one-third of total molar area, irrespective of molar proportions ($M2/(M1 + M2 + M3) = (a/i)/[1 + a/i + (2a/i - 1)] = 1/3$; see Methods). This is noteworthy because previous studies have found this relationship in primates⁴⁵, suggesting that the inhibitory cascade may be expected to apply across mammalian orders.

Even though we have shown here how the inhibitory cascade can be used to account for the evolutionary diversity of molar proportions, ecological and functional factors are still likely to have an indirect function in these differences. For example, previous analyses have shown that the overall crown complexity of rodent molars closely reflects the species-specific diets⁴². High crown-feature complexity is associated with herbivory, whereas simpler, smaller crowns are found in animal-eating taxa⁴². In our diversity data, the highly derived species with either specialized animal or fibrous-vegetation diets are plotted at the far ends of the molar-proportion spectrum (Fig. 5b). In other words, herbivorous murine species have more equal-sized teeth, whereas more faunivorous species, such as *Hydromys* (Fig. 5b), have progressively more reduced distal teeth. In comparison with dental complexity⁴², however, molar proportions seem not to be a measure of diet across mammalian orders because, for example, many herbivorous primates have progressively larger distal molars. We propose that molar proportions may not reflect function itself but may manifest the way in which development, by affecting the variational properties of teeth, responds to selection on functional features such as complexity and overall size.

Whereas our model predicts evolutionary change based on development, these predictions should not be taken as constraints on evolution. One clear exception is herbivorous arvicoline rodents (voles), in which the anterior part of their M1 is greatly elongated

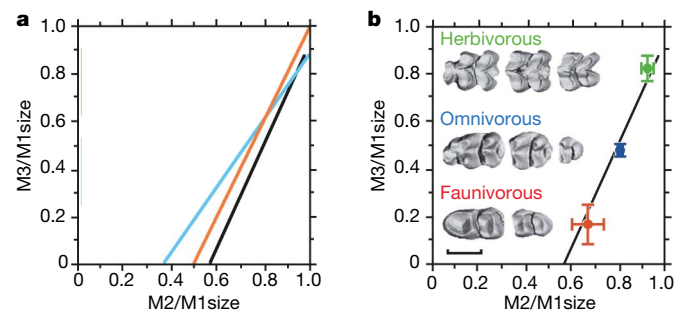


Figure 5 | The inhibitory cascade and the ecological context of murine dental diversity. **a**, From the macroevolutionary data (black line), the M2/M1 size was used to calculate the predicted M3/M1 size with the inhibitory cascade model (orange line; examples of molar proportions: $M1 = M2 = M3$; $M1 > M2 > M3$; $M1 \gg M2 \gg M3$). The random relay prediction illustrated (blue line: $M1 = M2 > M3$; $M1 > M2 = M3$; $M1 \gg M2 > M3$), for which randomized M2/M1 sizes were used to predict M3/M1 sizes, is the mean of reduced major axis regressions performed on each of 1,000 random simulations. All correlations, slopes and intercepts of the diversity data and the prediction of the inhibitory cascade model are significantly different from those of the 1,000 random relays ($P = 0.005$ to $P < 0.001$). **b**, The most equal molar proportions are found in herbivorous taxa and the least equal in faunivorous taxa, indicating that the inhibitory-cascade-influenced phenotypic change is under the control of ecology. The three examples of molar rows are scaled to body size (scale bar, 0.01 of body length) and are for *Mallomys rothschildi* (herbivore, $n = 2$), *Mus musculus* (omnivore, $n = 22$) and *Hydromys chrysogaster* (faunivore, $n = 3$), anterior towards the left. Error bars denote s.e.m. For details see Supplementary Information.

(refs 28, 46). This change can be considered a developmental novelty in which the M1 extension is allowed because there are no premolars in the anterior dental diastema. Nevertheless, we postulate a general situation in which any developmentally derived rule would predict that organisms should most often fall on the developmentally favoured evolutionary trajectory (Figs 4 and 5). In the inhibitory cascade, for example, many Old World primates and ungulates have a weak inhibitory cascade resulting in large distal molars ($M1 < M2 < M3$), but are still predicted to fall along the predicted trajectory. In contrast, other kinds of developmental change would be required for invasion of other parts of the morphospace. For example, evolving $M1 < M2 > M3$ proportions can be predicted to require a combination of low inhibition and specific early arrest of M3 development. However, for murine molars, the inhibitory cascade seems to have sufficed when murine rodents, the most taxonomically diverse mammalian group living, radiated into multiple adaptive zones.

Conclusions

The inhibitory cascade model is an activator–inhibitor network-derived model that allows the prediction of evolutionary paths in a given selective environment. These kinds of mechanistic model differ from classical correlation-based approaches (for example, genetic covariance) because the developmental mechanism is identified and there is greater conceptual continuity from genotype to phenotype³. To this end, the exact genetic underpinnings of the inhibitory cascade model remain to be identified. Whereas possible molecular level candidates include signalling molecules (and their inhibitors) such as BMPs, Activin A (Fig. 3) and Ectodysplasin⁴⁷, and transcription factors such as Pax9 (ref. 48), the inhibitory cascade may or may not be centred on the same genes in every species. Ultimately, with many more than 500 extant species and divergence times extending from the Pleistocene through to the middle Miocene⁴¹, murine rodents may provide excellent tests for the generality of high-level and low-level developmental rules. For this task, we would argue that the best tests of usefulness of identified developmentally derived rules are both the generality of the rule's use in other systems or taxa and the ability to demonstrate how development matters in explaining the evolution of phenotypes.

Because activator–inhibitor networks are a common mechanism in development, we suggest that inhibitory cascade-derived rules may apply in explaining the size relationships in adjacent organs beyond tooth development, particularly in other systems with sequentially developing organs or repeating elements. In insects, competition between developing body parts has been shown to affect the evolution of morphology^{2,5}, and the inhibitory cascade may also be understood as a form of sequential competition between adjacent organs. In teeth, our model resolves long-standing debates about the regulation of individual molar initiation and size, highlighting the essential role of inhibitors in shaping the entire dental system. Furthermore, our strategy of using the experimentally defined logic of organ systems to develop high-level testable models for predicting morphological evolution provides a blueprint for further exploration of evolutionary predictability in natural systems.

METHODS SUMMARY

Lower molar tooth germs were dissected from heterozygous *ShhGFP* mouse embryos³³ at day 13 or 14 after fertilization, as described previously^{30,34}. Posterior tails (giving rise to M2 and M3) of developing M1s were separated, and both pieces were cultured a short distance from each other. Explants were photographed daily from initiation to day 12 of culture, and the molar initiation date was determined on the basis of the first visible epifluorescence marking the formation of each primary enamel knot. Recombinant protein beads^{30,34} were placed on the posterior end of day 14 tooth germs. Tooth rows of 29 murine rodent species, representing the range of diets across the phylogeny within the subfamily, were scanned with a laser scanner and entered into the MorphoBrowser database (<http://morphobrowser.biocenter.helsinki.fi/>) as described previously⁴². Two-dimensional crown areas were measured from

images of 61 explants and scans of 29 species. For all measures, statistical differences between groups were tested by using Mann–Whitney *U*-tests, each with two-tailed exact significance levels, performed in SPSS version 11.0 (SPSS Inc.). Model randomizations and calculations of reduced major-axis regressions⁴⁹ were performed in a custom Visual Basic 6.0 program (Microsoft Corp.) and additional calculations of reduced major-axis regressions were performed in PAST (<http://folk.uio.no/ohammer/past/index.html>) (see Supplementary Information).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 April; accepted 7 August 2007.

- Alberch, P. & Gale, E. A. A developmental analysis of an evolutionary trend: Digital reduction in amphibians. *Evol. Int. J. Org. Evol.* **39**, 8–23 (1985).
- Nijhout, H. F. & Emlen, D. J. Competition among body parts in the development and evolution of insect morphology. *Proc. Natl Acad. Sci. USA* **95**, 3685–3689 (1998).
- Wagner, G. P., Chiu, C.-H. & Laubichler, M. Developmental evolution as a mechanistic science: the inference from developmental mechanism to evolutionary processes. *Am. Zool.* **40**, 819–831 (2000).
- Salazar-Ciudad, I. & Jernvall, J. How different types of pattern formation mechanisms affect the evolution of form and development. *Evol. Dev.* **6**, 6–16 (2004).
- Emlen, D. J., Hunt, J. & Simmons, L. W. Evolution of sexual dimorphism and male dimorphism in the expression of beetle horns: phylogenetic evidence for modularity, evolutionary lability, and constraint. *Am. Nat.* **166**, 42–68 (2005).
- Polly, P. D. Development and phenotypic correlations: the evolution of tooth shape in *Sorex araneus*. *Evol. Dev.* **7**, 29–41 (2005).
- Brakefield, P. M. & Roskam, J. C. Exploring evolutionary constraints in a task for an integrative evolutionary biology. *Am. Nat.* **168**, 4–13 (2006).
- Colosimo, P. F. *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).
- Protas, M. E. *et al.* Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genet.* **38**, 107–111 (2006).
- Prud'homme, B. *et al.* Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. *Nature* **440**, 1050–1053 (2006).
- Shapiro, M. D., Bell, M. A. & Kingsley, D. M. Parallel genetic origins of pelvic reduction in vertebrates. *Proc. Natl Acad. Sci. USA* **103**, 13753–13758 (2006).
- True, J. & Haag, E. S. Developmental system drift and flexibility in evolutionary trajectories. *Evol. Dev.* **3**, 109–119 (2001).
- Abouheif, E. & Wray, G. A. Evolution of the gene network underlying wing polymorphism in ants. *Science* **297**, 249–252 (2002).
- Kawasaki, K., Suzuki, T. & Weiss, K. M. Phenogenetic drift in evolution: the changing genetic basis of vertebrate teeth. *Proc. Natl Acad. Sci. USA* **102**, 18063–18068 (2005).
- Tanaka, M. *et al.* Developmental genetic basis for the evolution of pelvic fin loss in the pufferfish *Takifugu rubripes*. *Dev. Biol.* **281**, 227–239 (2005).
- Bateson, W. *Materials for the Study of Variation, Treated with Special Regard to Discontinuity in the Origin of Species* (Macmillan, London, 1894).
- Butler, P. M. Studies of the mammalian dentition. Differentiation of the post-canine dentition. *Proc. Zool. Soc. London (B)* **109**, 1–36 (1939).
- Kurtén, B. On the variation and population dynamics of fossil and recent mammal populations. *Acta Zool. Fenn.* **76**, 1–122 (1953).
- Van Valen, L. Growth fields in the dentition of *Peromyscus*. *Evol. Int. J. Org. Evol.* **16**, 272–277 (1962).
- Gould, S. J. & Garwood, R. A. Levels of integration in mammalian dentitions: an analysis of correlations in *Nesophontes micrus* (Insectivora) and *Oryzomys couesi* (Rodentia). *Evol. Int. J. Org. Evol.* **23**, 276–300 (1969).
- Sofaer, J. A., Bailit, H. L. & MacLean, C. J. A developmental basis for differential tooth reduction during Hominid evolution. *Evol. Int. J. Org. Evol.* **25**, 509–517 (1971).
- Osborn, J. W. in *Development, Function and Evolution of Teeth* (eds Butler, P. M. & Joysey, K. A.) 171–201 (Academic, London, 1978).
- Smith, B. H. Dental development and the evolution of life-history in Hominidae. *Am. J. Phys. Anthropol.* **8**, 157–174 (1991).
- Godfrey, L. R., Samonds, K. E., Jungers, W. L. & Sutherland, M. R. in *Primate Life Histories and Socioecology* (eds Kappeler, P. M. & Pereira, M. E.) 177–203 (Univ. of Chicago Press, Chicago, 2003).
- Boughner, J. C. & Dean, M. C. Does space in the jaw influence the timing of molar crown initiation? A model using baboons (*Papio anubis*) and great apes (*Pan troglodytes*, *Pan paniscus*). *J. Hum. Evol.* **46**, 253–275 (2004).
- Macchiarelli, R. *et al.* How Neanderthal molar teeth grew. *Nature* **444**, 748–751 (2006).
- Silvestri, A. R. & Singh, I. The unresolved problem of the third molar: Would people be better off without it? *J. Am. Dent. Assoc.* **134**, 450–455 (2003).
- Jernvall, J., Keränen, S. V. E. & Thesleff, I. Evolutionary modification of development in mammalian teeth: Quantifying gene expression patterns and topography. *Proc. Natl Acad. Sci. USA* **97**, 14444–14448 (2000).

29. Salazar-Ciudad, I. & Jernvall, J. A gene network model accounting for development and evolution of mammalian teeth. *Proc. Natl Acad. Sci. USA* **99**, 8116–8120 (2002).
30. Kassai, Y. *et al.* Regulation of mammalian tooth cusp patterning by Ectodin. *Science* **309**, 2067–2070 (2005).
31. Gaunt, W. A. An analysis of the growth of the cheek teeth of the mouse. *Acta Anat.* **54**, 220–259 (1963).
32. Gritli-Linde, A. *et al.* Shh signaling within the dental epithelium is necessary for cell proliferation, growth and polarization. *Development* **129**, 5323–5337 (2002).
33. Harfe, B. D. *et al.* Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell* **118**, 517–528 (2004).
34. Sahlberg, C., Mustonen, T. & Thesleff, I. Explant cultures of embryonic epithelium: Analysis of mesenchymal signals. *Methods Mol. Biol.* **188**, 373–382 (2002).
35. Åberg, T., Wozney, J. & Thesleff, I. Expression patterns of bone morphogenetic proteins (Bmps) in the developing mouse tooth suggest roles in morphogenesis and cell differentiation. *Dev. Dyn.* **210**, 383–396 (1997).
36. Wang, X. P. *et al.* Modulation of activin/bone morphogenetic protein signaling by follistatin is required for the morphogenesis of mouse molar teeth. *Dev. Dyn.* **231**, 98–108 (2004).
37. Klein, O. D. *et al.* Sprouty genes control diastema tooth development via bidirectional antagonism of epithelial–mesenchymal FGF signaling. *Dev. Cell* **11**, 181–190 (2006).
38. Ferguson, C. A. *et al.* Activin is an essential early mesenchymal signal in tooth development that is required for patterning of the murine dentition. *Genes Dev.* **12**, 2636–2649 (1998).
39. Jernvall, J., Åberg, T., Kettunen, P., Keränen, S. & Thesleff, I. The life history of an embryonic signaling center: BMP-4 induces p21 and is associated with apoptosis in the mouse tooth enamel knot. *Development* **125**, 161–169 (1998).
40. Plikus, M. V. *et al.* Morphoregulation of teeth: modulating the number, size, shape and differentiation by tuning Bmp activity. *Evol. Dev.* **7**, 440–457 (2005).
41. Jansa, S. A., Barker, F. K. & Heaney, L. R. The pattern and timing of diversification of Philippine endemic rodents: Evidence from mitochondrial and nuclear gene sequences. *Syst. Biol.* **55**, 73–88 (2006).
42. Evans, A. R., Wilson, G. P., Fortelius, M. & Jernvall, J. High-level similarity of dentitions in carnivores and rodents. *Nature* **445**, 78–81 (2007).
43. Garn, S. M., Lewis, A. B. & Kerewsky, R. S. Third molar agenesis and size reduction of the remaining teeth. *Nature* **200**, 488–489 (1963).
44. Polly, P. D. Variability in mammalian dentitions: size-related bias in the coefficient of variation. *Biol. J. Linn. Soc.* **64**, 83–99 (1998).
45. Lucas, P. W., Corlett, R. T. & Luke, D. A. Sexual dimorphism of tooth size in anthropoids. *Hum. Evol.* **1**, 23–29 (1986).
46. Guthrie, R. D. Variability in characters undergoing rapid evolution, an analysis of *Microtus* molars. *Evol. Int. J. Org. Evol.* **19**, 214–233 (1965).
47. Kangas, A. T., Evans, A. R., Thesleff, I. & Jernvall, J. Nonindependence of mammalian dental characters. *Nature* **432**, 211–214 (2004).
48. Kist, R. *et al.* Reduction of *Pax9* gene dosage in an allelic series of mouse mutants causes hypodontia and oligodontia. *Hum. Mol. Genet.* **14**, 3605–3617 (2005).
49. Sokal, R. R. & Rohlf, F. J. *Biometry* (Freeman, New York, 1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. K. Chapple, G. Evans, M. Fortelius, I. Salazar-Ciudad, M. Mikkola, I. Thesleff, G. P. Wilson and P. C. Wright for comments, discussions and support with this work; P. Munne, M. Mäkinen, E. Penttilä, I. Pljusnin, R. Santalahti and R. Savolainen for technical help; M. Hyvönen for activin A recombinant protein; C. Tabin and A. Gritli-Linde for the *ShhGFP*Cre mice; and the following museum curators and collection managers for loans: O. Grönwall, R. Asher, M. Hildén and I. Hanski. This study was supported by the Academy of Finland.

Author Contributions K.D.K. and J.J. conceived the study; K.D.K. performed developmental experiments; A.R.E. acquired three-dimensional data; K.D.K., A.R.E. and J.J. performed quantitative analyses; A.R.E. and J.J. constructed the model; A.R.E. performed computer simulations; K.D.K., A.R.E. and J.J. wrote the paper; and J.J. coordinated the study.

Author Information The three-dimensional scans for this study are deposited in the MorphoBrowser database, at <http://morphobrowser.biocenter.helsinki.fi/>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.D.K. (kathryn_kavanagh@yahoo.com) or J.J. (jernvall@fastmail.fm).

METHODS

Tooth cultures. Lower M1 tooth germs were dissected from heterozygous *ShhGFP* mouse *Shh^{tm1(EGFP/Cre)}Cjrl/+* embryos³³ at day 13 or 14 after fertilization and cultured at 37 °C and 5% CO₂ with a Trowell-type organ culture as described previously^{30,34}. In brief, teeth were placed on 0.1 µm Nucleopore filter paper (Whatman) on a raised wire grid in a small Petri dish containing 2 ml of tissue medium (45% DMEM (Gibco), 45% F12/Glutamax (Gibco), 10% fetal bovine serum (PAA Laboratories GmbH) and 1% penicillin–streptomycin (10 U ml⁻¹; Gibco)). Medium was replaced every two to three days, and ascorbic acid (100 µg ml⁻¹) was added. The epifluorescence of *ShhGFP* teeth closely follow the patterns of *Shh* expression detected with *in situ* hybridization techniques²⁸.

***In vitro* experiments.** Tooth germs were separated from the jaw tissue, which, if left in place, would grow and stunt the development of teeth in culture. Posterior tails of tooth germs in culture were separated from the tooth germ with a 25-gauge needle, and both pieces were cultured a short distance (about 100 µm) from each other on the same filter paper. For a clean cut, the filter paper with the tooth germ was briefly placed on a glass Petri dish for the cutting, taking care to avoid desiccation, then returned to the grid over medium. For day 13 tooth germs, the tail was cut from the point at which the initial anterior broadening stopped, or one-quarter of the way from the end of the tail. Digital images of 80 explants were taken daily (except for 80 out of 1,040 cases) from initiation to day 12 of culture under a fluorescence microscope (Leica MZFLIII microscope and Olympus DP50 digital camera system at magnifications of ×3.2 and ×4.0, resulting in 0.44 and 0.55 pixel µm⁻¹ resolutions, respectively). Molar initiation date was tabulated by tracking backwards from the final M2 or M3 to the first visible epifluorescence marking the formation of the primary enamel knot.

Recombinant protein bead experiments were performed as described previously^{30,34}. Agarose beads (Affi-Gel-Blue beads, catalogue no. 153-7302; Bio-Rad) were washed three times in PBS, then soaked in Activin A (100 ng µl⁻¹)⁵⁰, BMP4 (100 ng µl⁻¹; R&D Systems) or BSA control (1 µg µl⁻¹; Sigma). Roughly 50 beads were soaked in 5 µl of 100 ng µl⁻¹ protein solution for 45 min at 37 °C and a bead was placed with fine forceps on the posterior end of the day 14 tooth germ.

Quantitative analyses of experimental and macroevolutionary data. We chose the day 12 culture point for morphological measurements because at this stage M1 has reached, and M2 is close to reaching, asymptotic size and because after this day teeth are often difficult to measure accurately *in vitro* because of superfluous tissue growth and differentiation. From digital images we measured the two-dimensional areas of developing tooth crowns. Even when teeth have rolled

onto their side this gives a reasonably consistent measure of size because cultured teeth have a tendency to flatten. However, explants in which M1 was pointing vertically, thus providing a considerable underestimate of its size relative to other molars, were excluded from measurements. The areas of 61 M1s, 48 M2s and 17 M3s were measured with NIH Image 1.63 and ImageJ (<http://rsb.info.nih.gov/ij/>). In addition, molar sizes were measured on alternate days of culture for 21 explants.

The 29 murine rodent species used in the macroevolutionary analysis were selected to represent the range of diets across the phylogeny within the subfamily and were determined from published literature sources described previously⁴². A tooth row of each species was scanned with a Nextec Hawk three-dimensional laser scanner and entered into the MorphoBrowser database at <http://morphobrowser.biocenter.helsinki.fi/> (ref. 42). Teeth were oriented manually to maximize crown–base projection and crowns were captured with the JavaView viewing utility in MorphoBrowser. Two-dimensional areas were measured with NIH Image 1.63 and ImageJ. All ratios were plotted with the use of non-transformed mm² areas.

Developmental models. We assumed a linear effect of the activator and inhibitor ratio on tooth proportions, namely $(a - i)/i = (a/i) - 1$. Other relationships (for example, $\log(a/i)$) would alter the amount by which teeth changed along the inhibitory cascade trajectory but not the trajectory itself. Solving the molar sizes from $1 + [(a - i)/i](x - 1)$ gives $M1 = 1$, $M2 = a/i$ and $M3 = 2a/i - 1$. Molar proportions (of all the molars) are $M1 = i/3a$, $M2 = 1/3$ and $M3 = (2a - i)/3a$. From these formulae, the relationship between the M2/M1 ratio and the M3/M1 ratio is $M3/M1 = 2(M2/M1) - 1$. Note that the constant 1/3 proportion of M2 is lost if the tooth row has four molars. Progressively larger posterior molars may still be initiated sequentially if their growth rates are correspondingly faster. Even though the model may apply to volumes (or numbers of cells), we present our measurement data with two-dimensional surface areas because we consider these more reliable and because they are commonly used in morphological research. We note, however, that transforming the two-dimensional areas to volumes does not change the pattern of results. In the random relay model, we randomly reshuffled the M2/M1 sizes (in effect, the strength of $(a - i)/i$) before determining M3/M1 sizes. A total of 1,000 randomizations were performed in a custom Visual Basic 6.0 program and calculations of reduced major axis regressions were performed as described⁴⁹.

50. Harrington, A. E. *et al.* Structural basis for the inhibition of activin signalling by follistatin. *EMBO J.* 25, 1035–1045 (2006).

Enzymatic capture of an extrahelical thymine in the search for uracil in DNA

Jared B. Parker¹, Mario A. Bianchet², Daniel J. Krosky¹, Joshua I. Friedman¹, L. Mario Amzel² & James T. Stivers¹

The enzyme uracil DNA glycosylase (UNG) excises unwanted uracil bases in the genome using an extrahelical base recognition mechanism. Efficient removal of uracil is essential for prevention of C-to-T transition mutations arising from cytosine deamination, cytotoxic U•A pairs arising from incorporation of dUTP in DNA, and for increasing immunoglobulin gene diversity during the acquired immune response. A central event in all of these UNG-mediated processes is the singling out of rare U•A or U•G base pairs in a background of approximately 10^9 T•A or C•G base pairs in the human genome. Here we establish for the human and *Escherichia coli* enzymes that discrimination of thymine and uracil is initiated by thermally induced opening of T•A and U•A base pairs and not by active participation of the enzyme. Thus, base-pair dynamics has a critical role in the genome-wide search for uracil, and may be involved in initial damage recognition by other DNA repair glycosylases.

The viability of living organisms depends on efficient and highly specific enzymatic repair of the ubiquitous chemical damage that is inflicted on genomic DNA¹. One of the most conserved enzyme combatants in this battle to maintain genomic integrity is uracil DNA glycosylase (UNG), which locates undesirable uracil bases in DNA and then severs the bond between the base and deoxyribose sugar, initiating the process of uracil base excision repair². Efficient removal of uracil is essential for prevention of C→T transition mutations arising from cytosine deamination², cytotoxic U•A pairs arising from incorporation of dUTP in DNA³, and for increasing immunoglobulin gene diversity during the acquired immune response⁴.

One intriguing structural aspect of the recognition mechanism is that the enzyme interacts with uracil in DNA using an extrahelical recognition mechanism whereby the uracil base is expelled from the base stack, and engulfed deep in the active site pocket (FF state, Fig. 1)^{2,5,6}. The size of the uracil-specific active site sterically destabilizes binding of the larger thymine base, which differs from uracil by only a single methyl group at the 5-position of the pyrimidine ring. Consistent with this model, UNG has been converted into a thymine DNA glycosylase by a simple amino acid substitution that enlarges the active site⁷. The activity of such a mutant on a normal thymine base would, by necessity, require flipping of thymine from the DNA duplex, which raises the compelling question whether the search for uracil also involves flipping of normal thymine bases in a molecular quality control inspection process. Such a process was indeed supported by NMR dynamic measurements of T•A base-pair opening in the presence of UNG, where the enzyme was found to increase substantially the lifetime of the open state of thymine in T•A base pairs^{8,9}. In contrast, UNG had little effect on the base-pair opening rate as compared to the free DNA. These surprising findings suggested a passive role for the enzyme in trapping extrahelical pyrimidines, rather than actively accelerating their expulsion from the DNA duplex. The implication was that spontaneous thermally induced opening of T•A and U•A base pairs exposes an extrahelical state that sets in motion the remainder of the base-flipping process. Although these NMR studies suggest the presence of a weak

pyrimidine-binding site that is occupied transiently on the flipping reaction coordinate (EI, Fig. 1), the location and interactions of this site with the base are entirely unknown.

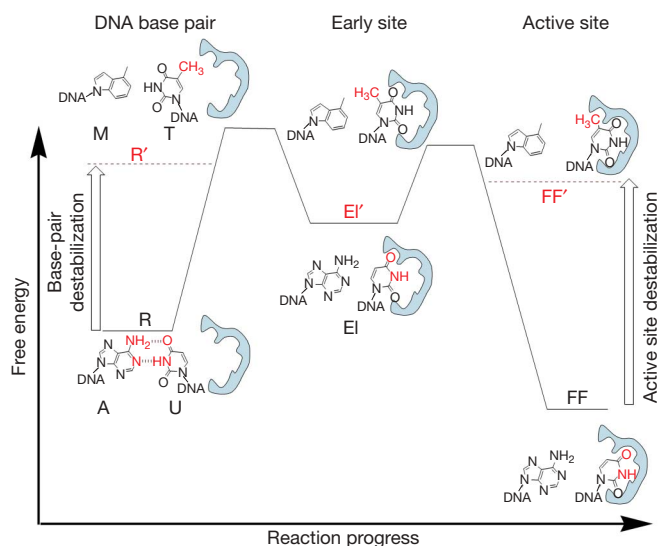


Figure 1 | Extrahelical uracil recognition by UNG and reaction coordinate tuning. Uracil emerges from the DNA base stack (reactant or R state) by spontaneous U•A base-pair breathing, where it is then trapped by UNG as an unstable early intermediate state (EI) on the base-flipping reaction coordinate. EI is very unstable (high energy), compared to the low-energy intrahelical bound state for a T•A or U•A base pair or the fully-flipped (FF) state^{8,9}. Substitution of adenine with its nonpolar analogue, 4-methylindole (M), energetically destabilizes the intrahelical R' state. Substitution of uracil with 5-methyluracil (5-MeU or thymine, T) greatly destabilizes the FF' state because the bulkier T fits poorly into the uracil active site, but T can access the EI' state^{8,9}. The energetic effects of reaction coordinate tuning on base flipping are shown by the vertical arrows. Destabilization of the R and FF states allows population of the otherwise unstable EI intermediate, allowing its structural characterization by X-ray crystallography. The free energy levels that are depicted in this figure are exaggerated for clarity of exposition.

¹Departments of Pharmacology and Molecular Sciences and ²Biophysics and Biophysical Chemistry of the Johns Hopkins Medical School, 725 North Wolfe Street, Baltimore, Maryland 21205, USA.

A cryptic pyrimidine-binding site on UNG

Structural characterization of transient reaction intermediates is a formidable challenge, made even more difficult in protein–DNA interactions by the occurrence of weak enzyme-binding modes that thwart formation of the unique complex of interest. Here we report the use of a reaction coordinate tuning method to populate a transitory state during enzymatic base flipping (Fig. 1). The approach is based on two principles obtained from extensive thermodynamic, NMR and rapid kinetic studies of base flipping by UNG^{8–11}. The first principle is that the equilibrium constant for base flipping may be shifted away from the reactant state by destabilization of the uracil base pair. Accordingly, the nonpolar adenine isostere 4-methylindole (M) possesses no hydrogen-bonding groups (Fig. 1)¹², and when placed opposite to uracil in DNA allows UNG to bind 8,000 times more tightly as compared with an identical DNA with a U•A base pair¹⁰. The second principle is that the fully extrahelical product state (FF, Fig. 1) can be destabilized by replacing the 5-hydrogen of uracil with a bulky methyl group to make 5-methyluracil (5-MeU or T) that sterically prevents access to the uracil pocket (Fig. 1)⁷. Combining these two principles flattens the energy landscape for base flipping and allows population of any transient extrahelical states that may exist between the reactant and product states (EI', Fig. 1).

The structure of human UNG bound to DNA containing a central T•M base pair refined to 2.5 Å resolution reveals that the central thymine is rotated from the base stack by about 30° (Fig. 2a and Supplementary Fig. 2a), which is only one-sixth of the 180° rotation required to fully flip uracil into the active site pocket (Protein Data Bank code 1EMH, Fig. 2b)⁶. Hence this structure represents a

structural snapshot of a very early intermediate on the base-flipping pathway. Despite the large differences in the positions of the extrahelical bases in the early intermediate (EI') and the previously reported fully flipped (FF) structure, the complexes share many of the same DNA backbone interactions, indicating that the initial steps in flipping uracil and thymine are identical (Fig. 2c). Most notably, a key DNA intercalating residue (Leu 272) protrudes into the minor groove of the DNA in both complexes (side chain shown in yellow in Fig. 2d), with its δCH_3 group lodged within van der Waals contact distance of the deoxyribose ring of the flipped nucleotide. This leucine has been shown in rapid kinetic studies to be important in promoting forward migration along the base-flipping pathway, with functional roles both early and late in the process¹³. The intercalated position of Leu 272 in the EI' complex is fully consistent with its proposed early role in flipping that involves plugging the gap in the base stack left behind by the extruded base, thereby inhibiting its retrograde motion back into the duplex. Additional interactions common to the EI' and FF complexes involve the phosphodiester groups flanking the extrahelical nucleotide, as shown in Fig. 2c for the EI' complex. These interactions include neutral hydrogen bonds from either the backbone amide or side-chain hydroxyl groups of residues Ser 169, Ser 270, Ser 273 and Ser 247 to the bridging or non-bridging phosphate oxygens (Supplementary Fig. 2a). The early emergence of these serine–phosphate interactions along the flipping pathway is consistent with previous kinetic and mutational studies¹³.

We conclude that T and U follow the same flipping reaction coordinate on the basis of the following lines of evidence. First, both bases emerge spontaneously from the duplex with similar rates kinetically

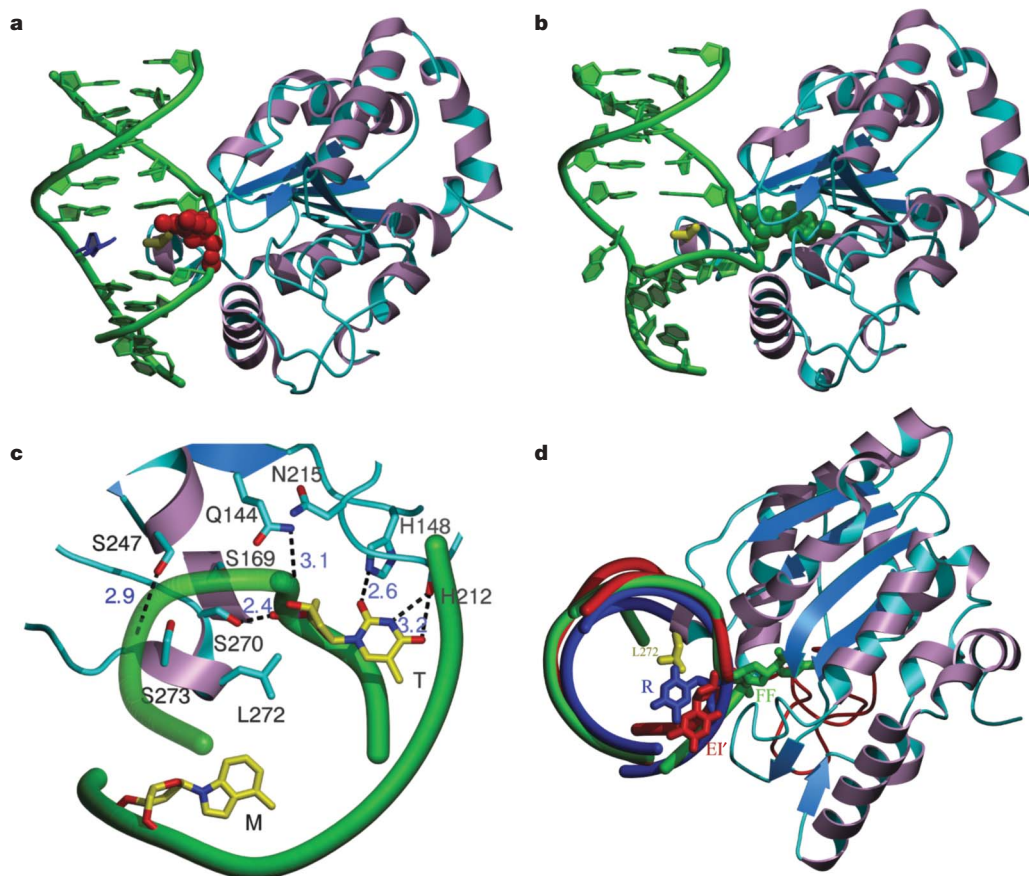


Figure 2 | Stabilization of extrahelical thymine in the EI' state in complex with UNG. **a**, View of the extrahelical thymine trapped in the EI' state with UNG. **b**, View of extrahelical uracil containing DNA in the FF state with UNG (Protein Data Bank code 1EMH)⁶. **c**, Interactions with the M•T base pair in the EI' state. Distance (in ångströms) is indicated by dotted lines.

d, Computational model of the R (reactant state) complex with an intrahelical thymine (blue) is shown aligned with the crystallographic models for the EI' (red) and FF (green) complexes. The view is rotated 90° out of the plane of the page relative to panels **a** and **b**. The extended coil region containing His 148 is indicated in red.

competent for base flipping (see further discussion below)^{8,9}. Second, the enzyme–DNA backbone interactions of T in the EI state are shared with those of U in the final FF state (see above). It is thus quite easy to see how these early interactions would lead to the FF state. Third, there are no interactions of the enzyme with the substituent at the 5 position of the base in the EI state. Thus, there is no obvious way that UNG could discriminate between the 5-CH₃ of T and the 5-H of U at this stage, strongly indicating that U and T occupy this same transient site (Fig. 2c). Finally, we have also obtained a crystal structure of the abasic product DNA arising from slow excision of thymine in the crystal over several weeks (Supplementary Fig. 2b). Thus, T is a very slow substrate that requires that it transiently accesses the active site. Therefore, the pathway for flipping T is productive, consistent with it following the same pathway as U.

Despite similar DNA backbone interactions in both complexes, the overall DNA structures are quite different for the EI' and FF complexes (Fig. 2a, b). For both structures, the DNA resembles B-form DNA 3' of the flipped nucleotide, with an average rise of ~10.5 base pairs per turn and standard 2'-endo sugar pucker. However, there is a ~20° shift in the helical axis immediately 5' to the flipped uracil in the FF complex that arises from the extreme conformation of the extrahelical nucleotide. In contrast, the EI' complex deviates in only a minor way from docked B-form DNA except for local perturbations at the immediate site of thymine extrusion (compare Fig. 2a and b). The absence of extreme DNA bending in the EI' complex is consistent with fluorescence and NMR studies of the early stages of base flipping that indicated little perturbation of the DNA base stack and preservation of the B-form DNA conformation⁸. This aspect of the UNG reaction differs from observations obtained from structural studies of DNA complexes with human 8-oxoguanine DNA glycosylase (hOGG1), in which the enzyme seems to bend DNA regardless of whether the cognate oxidized base (8-oxoguanine) is present^{14,15}. DNA bending is a key component of the base extrusion process because it allows widening of the major groove, providing an egress pathway for the base, and also serves to release some of the strain resulting from the DNA backbone distortions that accompany flipping.

The most marked difference between the EI' and FF complexes is the position of the extrahelical thymine and uracil base (Fig. 2a–d). The thymine base, which is highly exposed to solvent, docks against two regions of UNG that form the mouth of the active site. A key interaction is a charged hydrogen bond ($d = 2.6 \text{ \AA}$) that is formed between thymine O2 and the side-chain NH ϵ proton of the completely conserved His 148, which is located near the beginning of a long coiled region of the protein backbone that stretches from residues Gln 144 to Pro 167 (thin red strand, Fig. 2d). A second interaction is observed between the imino proton of thymine and the backbone carbonyl of His 212 ($d = 3.2 \text{ \AA}$), which is located in a highly conserved nine-residue turn that encompasses residues Ala 211 to Glu 219. Aside from these two hydrogen bonds, there are no other interactions of UNG with the thymine base. Nevertheless, these limited interactions could provide specificity for uracil and thymine because neutral cytosine is not complementary with this hydrogen bond donor–acceptor pattern. In addition, bulky purines would be sterically excluded from the site owing to the tight packing of Leu 272 against the deoxyribose, which fixes the position of the deoxyribose–base glycosidic bond. In contrast with the relatively sparse interactions observed in EI', the uracil base is extensively desolvated in the FF complex, with every potential hydrogen bond donor and acceptor interaction fulfilled, and in addition, favourable edge–face aromatic interaction between Phe 158 and the uracil ring⁶. The increasing interactions with the base and the phosphate backbone as the reaction proceeds indicate that base flipping occurs within an enzyme energy landscape that serves to drive the reaction forward in a succession of energetically downhill steps^{10,13}.

Conformational flexibility in flipping

The EI structure reveals that the flipped nucleotide undergoes unexpected conformational changes over the reaction coordinate (Fig. 3). In the EI state, the plane of the thymidine sugar has rotated by about 30° relative to the B-DNA reactant state, but the thymine base has rotated around the glycosidic torsion angle (χ) by 180° to place it in an unusual *syn* conformation and 3'-exo sugar pucker. This glycosidic bond rotation would nicely serve to present the Watson–Crick edge of the base to the enzyme binding pocket, allowing it to recognize specifically the hydrogen bond donor–acceptor pattern of uracil and thymine. Finally, migration of uracil into the active site requires a further 120° rotation of the sugar plane and, also, another 90° rotation of the base around χ (Fig. 3). The unexpected conformation of thymine in EI suggests that rapid, unrestrained rotation around χ occurs at one or more steps along the reaction path to the active site.

The capture of thymine in the transient pyrimidine binding site on UNG corresponds to an earlier stage of base flipping than a previous structure of hOGG1, in which an extrahelical guanine was trapped using disulphide crosslinking technology¹⁴. In the hOGG1 structure, the sugar of the flipped guanine nucleotide was rotated nearly 150°, which is nearly the full rotation required to move it from the DNA base stack into the active site. In contrast, the thymine sugar is only rotated about 30° when it occupies the transient binding site on UNG. In a similar manner to UNG, hOGG1 also uses an imidazole NH (His 270) to stabilize the flipped guanine in its binding site. In both enzymes, the imidazole–base interactions move to the phosphate backbone in the final pre-catalytic complex.

Base-pair-opening dynamics

As summarized in Fig. 4a, the structure of the EI' complex reveals discrete interactions that, when deleted, should have a deleterious impact on UNG's ability to stabilize the extrahelical thymine. To explore the energetic role of these groups, we mutated residues corresponding to His 148, Gln 144, Leu 272, Ser 247, Ser 169 and Ser 270 to alanine or glycine and then measured the effects on the NMR imino proton exchange rates with solvent using a symmetric 10-mer duplex with a central T•A base pair (Fig. 4a and Supplementary Figs 3 and 4)^{16,17}. Previous NMR studies have established that UNG accelerates the observed imino proton exchange rate of the central thymine by 25 times that of the same proton in the free DNA duplex⁸, and that these effects on spontaneous thymine base-pair opening are not unique to this DNA sequence⁹. The observed acceleration arises from a 75-fold increase in the equilibrium concentration of the solvent-exposed open state from which exchange occurs ($K_{\text{open}} = k_{\text{open}}/k_{\text{close}}$). The increase in K_{open} arises

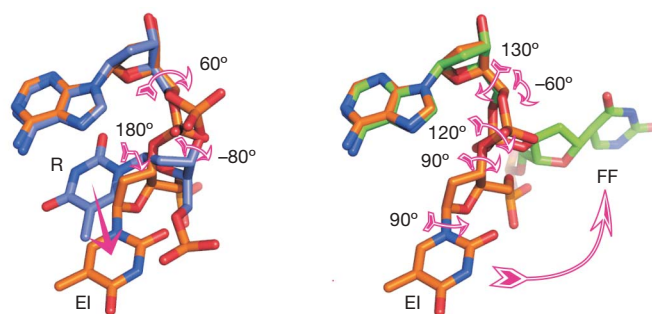


Figure 3 | Conformational changes in the sugar and base along the flipping reaction coordinate. In the first step of the reaction to form EI, the sugar plane rotates about an apparent angle of 30°, and the base rotates 180° around the glycosidic bond and moves about 4.4 Å relative to the B-DNA reactant state (R, blue). In the second step (EI→FF), the sugar plane and base rotate a further 120° and 90°, respectively. Note that the structure of the FF state was obtained using the C-glycoside analogue of deoxyuridine, pseudodeoxyuridine. Changes in the DNA backbone torsional angles that accompany these transformations are listed in the figure.

almost entirely from an enzyme-induced decrease in the closing rate (k_{close}) and not an acceleration of the spontaneous opening rate (k_{open}) as compared to the free DNA^{8,9}.

The imino proton exchange time courses for the central thymine in the presence of wild-type UNG and the mutants are shown in Fig. 4b (see also Supplementary Tables 2 and 3, and Supplementary Figs 3 and 4). In general, the mutations reduced the exchange rate in the range of twofold to tenfold (Fig. 4c), confirming that these interactions stabilize the EI state. Unlike wild-type UNG, we were unable to detect general base catalysis of imino proton exchange with any of

the exchange-defective mutants. This finding indicates that, for these mutant enzymes, the imino position is inaccessible to the difluoroethylamine base catalyst. In contrast, its exposed position observed in the complex with wild-type UNG (Fig. 2c) facilitated robust general base catalysis and allowed measurement of the microscopic rate and equilibrium constants for the two-step exchange process (Fig. 4b)⁸. The largest effects observed with the L272G, S169A and S270A mutant enzymes decrease the exchange rate of the central thymine by tenfold, to nearly the level observed in the free DNA duplex (Fig. 4c). The large effect on the exchange rate on removal of Ser 169 may be indirect because its O γ is located over 4 Å from the phosphate (Fig. 4a). However, Ser 169 also contacts O γ of Ser 270, which shows a tenfold reduction on the exchange rate, and Ser 270 is positioned only 2.4 Å from the 5' phosphate of the flipped thymine. These mutational effects on imino exchange are similar to the 39-, 5- and 10-fold effects of these same mutations on the stability of the FF state¹³, supporting the contention that these interactions are relevant early and late on in the reaction coordinate for thymine and uracil. With respect to stabilization of the EI state, it seems that the leucine plug and the serine–DNA phosphate interactions are more important than the His 148 hydrogen bond to O2 of the base. Surprisingly, the Q144A mutation did not result in slowing of the exchange rate even though the Gln 144 side-chain amide forms a 3.1 Å hydrogen bond to the phosphate backbone (Fig. 4a, c). By modelling, we find that removal of Gln 144 accommodates the change to the most common side-chain rotamer of Asn 215 that moves its amide to within 3.2 Å of the target phosphate (Fig. 2c), providing a plausible structural basis for the absence of a mutational effect.

These results highlight the importance of using complementary methods to distinguish between active and passive mechanisms for base flipping, because static structures are incapable of revealing the order of events. In this regard, previous structural studies of the DNA complexes of hOGG1 and MutM did not address the issue of whether these enzymes actively flip 8-oxoguanine from the duplex or, on the other hand, use a passive trapping mechanism as observed here for UNG^{14,18}. For UNG-bound DNA, the initial event of spontaneous base-pair opening may be estimated to occur with a rate constant $k_{\text{open}} = 1,400 \text{ s}^{-1}$ at 25 °C, based on our measured opening rate at 10 °C and previously measured activation enthalpies for base-pair opening (16–19 kcal mol⁻¹)^{19,20}. This rate is easily fast enough to account for the observed rates of enzymatic uracil flipping at 25 °C (<700 s⁻¹)¹³.

This study illustrates the significance of DNA base-pair dynamics in damage base recognition, and provides an interesting example of enzymatic recognition of a transient high energy conformation of a substrate, rather than the ground state. Notably, UNG does not distinguish uracil from thymine while it lies within the DNA helix, as previously proposed for 8-oxoguanine recognition by MutM glycosylase¹⁸, but instead discrimination occurs in a transient pyrimidine sieving pocket on the enzyme. This pocket solves the central kinetic problem in efficient base flipping: the extremely rapid re-entry of the base into the DNA duplex. Hence, by slowing down the retrograde motion, UNG promotes efficient and selective forward commitment of uracil down the remainder of the flipping pathway. Finally, the reaction coordinate tuning method used here offers an attractive and potentially less perturbing alternative to the disulphide crosslinking approach for trapping complexes of enzymes with unstable extrahelical bases²¹.

METHODS SUMMARY

The special 4-methylindole (M) phosphoramidite was synthesized according to published procedures^{12,22}. The DNA strands used for obtaining the crystal structure of the UNG–EI' complex were synthesized using standard solid-phase DNA synthesis methods from commercially available nucleoside phosphoramidites and hybridized to form the T•M duplex (strand 1, 5'-TGTTATCTT-3'; strand 2, 5'-AAAGATMACA-3'). Purification of bacterial and human UNG has been

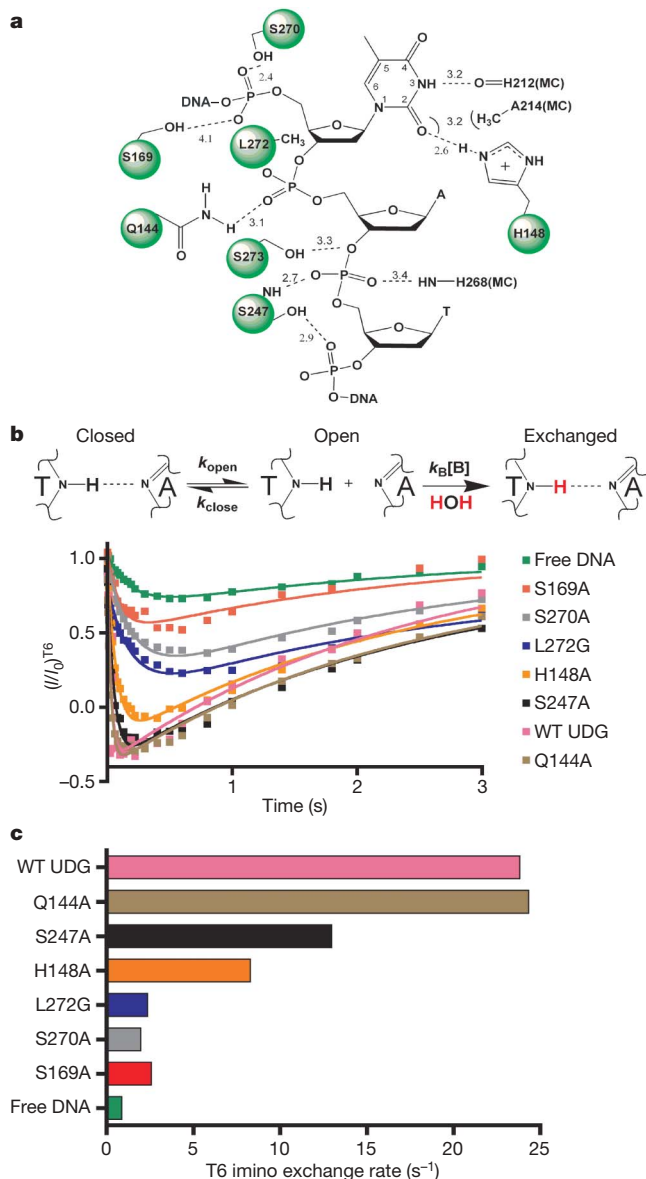


Figure 4 | Interaction map for the EI' complex and imino proton exchange profiles for wild-type UNG and several mutant forms. a, Side-chain and backbone interactions of UNG with the extrahelical thymine base and DNA phosphate backbone in the EI' complex (MC, main chain). **b**, Imino proton exchange is a two-step process that requires base-pair opening ($K_{\text{open}} = k_{\text{open}}/k_{\text{close}}$) followed by imino proton exchange with solvent, which may be followed by using NMR magnetization transfer from water¹⁶. The data are fitted to equation (1) (see Methods) and the exchange time courses for wild-type UNG and six mutants are indicated. $k_{\text{B}}[\text{B}]$ is the buffer (B)-catalysed exchange rate, and $(I/I_0)^{T_6}$ is the ratio of the T6 imino proton resonance intensity in the presence (I) and absence (I_0) of magnetization transfer from solvent. **c**, Exchange rates (k_{ex}) for T6 in the presence of wild-type UNG and the indicated mutants.

described^{23–26}. NMR magnetization transfer experiments were performed as described previously^{8,9}.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 March; accepted 27 July 2007.

Published online 19 August 2007.

- Lindahl, T. & Wood, R. D. Quality control by DNA repair. *Science* **286**, 1897–1905 (1999).
- Stivers, J. T. & Jiang, Y. L. A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem. Rev.* **103**, 2729–2759 (2003).
- Seiple, L., Jaruga, P., Dizdaroğlu, M. & Stivers, J. T. Linking uracil base excision repair and 5-fluorouracil toxicity in yeast. *Nucleic Acids Res.* **34**, 140–151 (2006).
- Kavli, B., Otterlei, M., Slupphaug, G. & Krokan, H. E. Uracil in DNA—General mutagen, but normal intermediate in acquired immunity. *DNA Repair* **6**, 505–516 (2006).
- Slupphaug, G. *et al.* A nucleotide-flipping mechanism from the structure of human uracil–DNA glycosylase bound to DNA. *Nature* **384**, 87–92 (1996).
- Parikh, S. S. *et al.* Uracil–DNA glycosylase–DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl Acad. Sci. USA* **97**, 5083–5088 (2000).
- Kavli, B. *et al.* Excision of cytosine and thymine from DNA by mutants of human uracil–DNA glycosylase. *EMBO J.* **15**, 3442–3447 (1996).
- Cao, C., Jiang, Y. L., Stivers, J. T. & Song, F. Dynamic opening of DNA during the enzymatic search for a damaged base. *Nature Struct. Mol. Biol.* **11**, 1230–1236 (2004).
- Cao, C., Jiang, Y. L., Krosky, D. J. & Stivers, J. T. The catalytic power of uracil DNA glycosylase in the opening of thymine base pairs. *J. Am. Chem. Soc.* **128**, 13034–13035 (2006).
- Krosky, D. J., Song, F. & Stivers, J. T. The origins of high-affinity enzyme binding to an extrahelical DNA base. *Biochemistry* **44**, 5949–5959 (2005).
- Krosky, D. J., Schwarz, F. P. & Stivers, J. T. Linear free energy correlations for enzymatic base flipping: How do damaged base pairs facilitate specific recognition? *Biochemistry* **43**, 4188–4195 (2004).
- Moran, S., Ren, R. X. F., Sheils, C. J., Rumney, S. & Kool, E. T. Non-hydrogen bonding ‘terminator’ nucleosides increase the 3′-end homogeneity of enzymatic RNA and DNA synthesis. *Nucleic Acids Res.* **24**, 2044–2052 (1996).
- Jiang, Y. L. & Stivers, J. T. Mutational analysis of the base flipping mechanism of uracil DNA glycosylase. *Biochemistry* **41**, 11236–11247 (2002).
- Banerjee, A., Yang, W., Karplus, M. & Verdine, G. L. Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* **434**, 612–618 (2005).
- Chen, L., Haushalter, K. A., Lieber, C. M. & Verdine, G. L. Direct visualization of a DNA glycosylase searching for damage. *Chem. Biol.* **9**, 345–350 (2002).
- Gueron, M. & Leroy, J. Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol.* **261**, 383–413 (1995).
- Snoussi, K. & Leroy, J. L. Alteration of A•T base-pair opening kinetics by the ammonium cation in DNA A-tracts. *Biochemistry* **41**, 12467–12474 (2002).
- Banerjee, A., Santos, W. L. & Verdine, G. L. Structure of a DNA glycosylase searching for lesions. *Science* **311**, 1153–1157 (2006).
- Gueron, M. *et al.* Applications of imino proton exchange to nucleic acid kinetics and structures. In *Structure and Methods* (ed. Sarma, R. H.) 113–137 (Adenine press, New York, 1990).
- Moe, J. G. & Russu, I. M. Kinetics and energetics of base-pair opening in 5′-d(CGCGAATTCGCG)-3′ and a substituted dodecamer containing G•T mismatches. *Biochemistry* **31**, 8421–8428 (1992).
- Verdine, G. L. & Norman, D. P. Covalent trapping of protein–DNA complexes. *Annu. Rev. Biochem.* **72**, 337–366 (2003).
- Ren, R. X.-F., Chaudhuri, N. C., Paris, P. L., Rumney, S. & Kool, E. T. Naphthalene, phenanthrene, and pyrene as DNA base analogues: Synthesis, structure, and fluorescence in DNA. *J. Am. Chem. Soc.* **118**, 7671–7678 (1996).
- Drohat, A. C., Jagadeesh, J., Ferguson, E. & Stivers, J. T. The role of electrophilic and base catalysis in the mechanism of *Escherichia coli* uracil DNA glycosylase. *Biochemistry* **38**, 11866–11875 (1999).
- Drohat, A. C. *et al.* Hetronuclear NMR and crystallographic studies of wild-type and H187Q *Escherichia coli* uracil DNA glycosylase: Electrophilic catalysis of uracil expulsion by a neutral histidine 187. *Biochemistry* **38**, 11876–11886 (1999).
- Xiao, G. *et al.* Crystal structure of *Escherichia coli* uracil DNA glycosylase and its complexes with uracil and glycerol: structure and glycosylase mechanism revisited. *Proteins* **35**, 13–24 (1999).
- Slupphaug, G. *et al.* Properties of a recombinant human uracil–DNA glycosylase from the *UNG* gene and evidence that *UNG* encodes the major uracil–DNA glycosylase. *Biochemistry* **34**, 128–138 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Majumdar for assistance with NMR experiments. This work was supported by NIH grants (J.T.S. and L.M.A.) and a major research instrumentation grant from the NSF.

Author Contributions J.B.P. prepared all mutant enzymes and performed the NMR experiments; M.A.B. collected, processed and refined X-ray data and performed structural analyses; D.J.K. conceived the structural approach and purified and crystallized the complexes; J.I.F. performed NMR studies on the L272G mutant; L.M.A. performed structural analyses, discussed and commented on the manuscript; J.T.S. analysed data and wrote the paper.

Author Information Coordinates and structure factor files for the EI complex and the abasic DNA complex have been deposited in the Protein Data Bank with accession numbers 2OXM and 2OYT, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.T.S. (jstivers@jhmi.edu).

METHODS

Materials. The 10-mer T•A duplex for the NMR studies (5'-C₁T₂G₃G₄A₅T₆C₇C₈A₉G₁₀-3') was purchased from Integrated DNA Technologies. It was purified by anion exchange HPLC (Zorbax) and desalted using a C-18 reversed phase HPLC column (Phenomenex Aqua column). The purity of all oligonucleotides was assessed by matrix-assisted laser desorption mass spectroscopy and denaturing polyacrylamide gel electrophoresis. The extinction coefficient of the duplex has been determined to be 144 mM⁻¹ cm⁻¹ at 260 nm⁸.

Purification of bacterial UNG and mutants. As previously described, recombinant UNG from *Escherichia coli* strain B was expressed using a T7 polymerase-based overexpression system and purified to >99% homogeneity^{23–25}. The concentration of the enzyme was determined using an extinction coefficient of 38.5 mM⁻¹ cm⁻¹. All of the mutations in this work were generated using the Quick-Change double-stranded mutagenesis kit from Stratagene, and the mutations were confirmed by sequencing both strands of the DNA. The purity of the mutant enzymes was greater than 95% as judged by SDS–polyacrylamide gel electrophoresis with visualization by Coomassie blue staining.

Crystallization of the complexes of human UNG and DNA. Human UNG2 was expressed and purified as described previously²⁶. A solution of human UNG2 (0.5 µl, 22 mg ml⁻¹) in a buffer containing 50 mM Tris-OAc, pH 7.0, 150 mM NaCl and 1 mM dithiothreitol (DTT), was mixed with T•M duplex DNA (0.5 µl, 2.5 mM). The mixture was allowed to incubate at ambient temperature for 30 min, and then centrifuged at 10,000g for 5 min. Co-crystallization condition was 22% PEG 4000, 100 mM sodium HEPES, pH 6.5 and 1 mM DTT. A total of 1 µl of the complex was mixed with an equal volume of precipitant, and allowed to crystallize at 22 °C using the hanging-drop method. Crystals were observed within 48 h. Crystals grew in tight agglomerates. Separation of a single crystal produced considerable stress in the crystals, which resulted in high mosaicity. A piece of a single crystal harvested a few days after growth, with a mosaicity of 2.6°, was used for the EI complex structure. A search to obtain isolated or more robust crystals was performed. Trials with 0.5% to 1% v/v of dioxane produced isolated crystals, between two to three weeks after set up. These high-quality crystals diffracted beyond 2.0 Å resolution in a home source; however, all crystals investigated were of a DNA duplex with an abasic residue, due to enzymatic activity during crystal growth (Supplementary Fig. 2b).

X-ray crystallography and structure determination. The two crystal forms in their unmodified mother liquor were flash frozen, and X-ray diffraction data were collected using an in house X-ray source. The package HKL2000 (ref. 27) was used for data reduction. The structures were determined by molecular replacement methods using the program PHASER²⁸. The protein portion of the complex of UNG bound to uracil-containing DNA (Protein Data Bank 1EMH) was used as a search model for both refinements. After initial rigid-body

refinement, difference electron density maps showed interpretable nucleotide density (Supplementary Fig. 2). The final model of the EI complex, refined using REFMAC5 (ref. 29) with isotropic temperature factors, has excellent stereochemistry and all but two non-glycine/proline residues in allowed regions of the Ramachandran plot (Asn 215 and Phe 155). The data collection and refinement statistics for both crystal forms are reported in Supplementary Table 1.

Docking of B-form DNA onto UNG. A model of B-form DNA (strand 1, 5'-TGTTATCTT-3'; strand 2, 5'-AAAGATAACA-3') was built using the program Quanta (Accelrys Inc.). An initial docking, performed using the PYMOL pair-fitting utility (Delano Scientifics LLC), was followed by manual adjustment. Figures were drawn using MOLSCRIPT and PYMOL³⁰, and rendering was performed with POVray 3.6.

NMR sample preparation. The 10-mer DNA duplex was hybridized in NMR buffer (1 mM Tris-d6·HCl, 35 mM NaCl, pH 8.0 at 10 °C). When preparing samples of the complex, 1 mM *E. coli* UNG and 0.8 mM 10-mer T•A DNA duplex were combined in a 500 µl total volume at high salt concentration (~500 mM NaCl) to avoid UNG precipitation, and then diluted with NMR buffer several times using a Microcon YM-3 centrifugal concentrator (Millipore) to provide a final NaCl concentration of 35 mM⁸. Data were acquired using 0.3-ml samples in 5-mm Shigemi reduced-volume NMR tubes.

NMR spectroscopy. NMR spectra were acquired using either Varian Unity Plus 500 MHz or Inova 600 MHz spectrometers (Varian NMR Systems) equipped with 5-mm triple resonance probes with z-axis-pulsed magnetic field gradients. Imino proton exchange rates were determined using NMR magnetization transfer from water using a gradient-enhanced pulse sequence as previously described⁸. The data were fitted to equation (1), where $I_z(t_{\text{mix}})$ is the intensity of an imino proton resonance at mixing time t , $I_{z,\text{eq}}$ is the equilibrium intensity, E is the efficiency of water inversion, k_{ex} is the exchange rate, and R_{li} and R_{lw} are the longitudinal relaxation rates for the imino proton and water, respectively.

$$\frac{I_z(t_{\text{mix}})}{I_{z,\text{eq}}} = 1 + Ek_{\text{ex}} \left[\frac{\exp(-R_{\text{li}} t_{\text{mix}}) - \exp(-R_{\text{lw}} t_{\text{mix}})}{R_{\text{lw}} - R_{\text{li}}} \right] \quad (1)$$

The exchange time courses and nonlinear regression fits to equation (1) for each imino proton of each DNA–enzyme complex are shown in Supplementary Fig. 5 and the exchange rates are reported in Supplementary Table 2.

- Otwinoski, Z. & Minor, W. Processing of X-ray diffraction data in oscillation mode. *Methods Enzymol.* **276**, 307–325 (1997).
- McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61**, 458–464 (2005).
- Collaborative Computational Project, Number 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–776 (1994).
- Esnouf, R. M. Further additions to Molscript version 1.4, including reading and contouring of electron density maps. *Acta Crystallogr. D* **55**, 938–940 (1999).

LETTERS

Coherent quantum state storage and transfer between two phase qubits via a resonant cavity

Mika A. Sillanpää¹, Jae I. Park¹ & Raymond W. Simmonds¹

As with classical information processing, a quantum information processor requires bits (qubits) that can be independently addressed and read out, long-term memory elements to store arbitrary quantum states^{1,2}, and the ability to transfer quantum information through a coherent communication bus accessible to a large number of qubits^{3,4}. Superconducting qubits made with scalable microfabrication techniques are a promising candidate for the realization of a large-scale quantum information processor^{5–9}. Although these systems have successfully passed tests of coherent coupling for up to four qubits^{10–13}, communication of individual quantum states between superconducting qubits via a quantum bus has not yet been realized. Here, we perform an experiment demonstrating the ability to coherently transfer quantum states between two superconducting Josephson phase qubits through a quantum bus. This quantum bus is a resonant cavity formed by an open-ended superconducting transmission line of length 7 mm. After preparing an initial quantum state with the first qubit, this quantum information is transferred and stored as a nonclassical photon state of the resonant cavity, then retrieved later by the second qubit connected to the opposite end of the cavity. Beyond simple state transfer, these results suggest that a high-quality-factor superconducting cavity could also function as a useful short-term memory element. The basic architecture presented here can be expanded, offering the possibility for the coherent interaction of a large number of superconducting qubits.

A particularly interesting quantum information architecture involves the interaction of atoms with optical or microwave cavities having discrete electromagnetic modes, or cavity quantum electrodynamics (QED). Cavity QED systems^{14,15} have enabled fundamental tests of quantum mechanics, as well as demonstrations of quantum memory and a quantum bus¹⁶. Recently, the Cooper pair box⁵ has been successfully incorporated into a superconducting resonant cavity in order to perform analogous experiments in the strong coupling regime, generating a new field of study known as circuit QED^{17–21}. Similar resonant cavities have been used to stabilize flux qubits²². Thus far, experiments have found spectroscopic evidence for the entanglement between two phase qubits and a resonator²³. In this circuit QED experiment, we report the first time-domain measurements showing coherent interactions between two phase qubits and a cavity formed by a transmission-line resonator. Moreover, by coupling two phase qubits to a single cavity, taking advantage of the independent control of each phase qubit and single-shot readout, we have constructed an elementary quantum memory and quantum bus in a superconducting system.

For a flux-biased Josephson phase qubit²⁴, the ground state $|g\rangle$ and the first excited state $|e\rangle$ are encoded in the phase difference δ across a large-capacitance superconducting Josephson junction placed in a superconducting loop (Fig. 1a). These states resemble those of a simple harmonic oscillator but for the nonlinear,

anharmonic potential²⁵ formed by the combination of the Josephson coupling energy $-E_J \cos(\delta)$ and the inductive energy stored in the superconducting loop, where E_J is the Josephson energy. Owing to their large capacitance, addressability, single-shot readout, and the ease with which the energy level separation $\hbar\omega = E_e - E_g$ can be tuned, phase qubits have proved to be relatively easy to couple together^{11,23}. Superconducting qubits offer the possibility of forming a quantum processor with the help of communication channels or a ‘qubus’.

Our superconducting quantum system is presented in Fig. 1b, c. Both qubits A and B are inductively coupled to two separate flux bias coils: one set of coils is used to adjust a static, direct-current (d.c.) flux bias, whereas the other set of radio-frequency coils, with a bandwidth from d.c. up to about 20 GHz, enables rapid flux bias changes (‘shift pulses’), inductively coupled microwave pulses, and a fast measurement pulse. Each set of qubit d.c. flux bias lines includes low-pass and copper powder filters, while each set of radio-frequency flux pulsed lines are combined into a single microwave coaxial line at room temperature and attenuated by roughly 40 dB inside the cryostat. Microwave pulse control is performed with passively filtered (roughly gaussian-shaped pulses) and standard microwave mixers. Independently addressable state readout is accomplished via inductively coupled d.c. superconducting quantum interference devices (SQUIDs).

Our resonant cavity is an open-ended coplanar waveguide whose lowest-standing-wave eigenmode ($\lambda/2$ -mode) has voltage maxima at each end of the waveguide (Fig. 1b). At near-resonance, this waveguide acts like a parallel, lumped element resonant LC circuit or cavity (Fig. 1c). The $\lambda/2$ -mode forms a simple harmonic oscillator with an energy $\hat{H}_r = \hbar\omega_r(\hat{a}^\dagger\hat{a} + \frac{1}{2})$ at the frequency $\omega_r/2\pi = 1/2\pi\sqrt{LC} \approx 8.74$ GHz, where $L = 2Z_0/\pi\omega_r$ and $C = \pi/2\omega_r Z_0$ represent their lumped element equivalents, $Z_0 \approx 50\Omega$ is the characteristic impedance of the coplanar waveguide, and the raising and lowering operators \hat{a}^\dagger and \hat{a} increase or decrease the photon number in the cavity.

Using the rotating-wave approximation, the hamiltonian of our quantum system formed by a single resonant cavity coupled at both ends to qubits A and B has the form of the Jaynes–Cummings hamiltonian^{14,15} familiar from quantum optics

$$\hat{H} = \hat{H}_r + \sum_{j=A,B} \hat{H}_j + \sum_{j=A,B} \hbar g_j (\hat{a}^\dagger \hat{\sigma}_-^j + \hat{a} \hat{\sigma}_+^j) \quad (1)$$

where $\hat{H}_j = \frac{1}{2}\hbar\omega_j \hat{\sigma}_+^j \hat{\sigma}_-^j$ is the single-qubit hamiltonian, $\hat{\sigma}_+^j$ ($\hat{\sigma}_-^j$) is the raising (or lowering) operator for creating (or annihilating) excitations in the j th qubit, and $\hbar\omega_j$ is controlled by the amplitude of the d.c. and radio-frequency flux bias. The interaction energy $2g_{A,B} \approx \omega_r (C_c / \sqrt{CC_{j,A,B}})$ was designed to be large enough to ensure that the timescale of quantum state transfer, $\pi/g_{A,B} \approx 10$ ns, would

¹National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA.

not be limited by the relaxation times of either qubit or the cavity, putting this experiment in the strong coupling regime ($g_{A,B} > \gamma_{A,B} > \kappa$) for circuit QED¹⁹ with qubit decay rates of $\gamma_A \approx 20$ MHz and $\gamma_B \approx 5$ MHz, and a cavity decay rate of $\kappa/2\pi \approx 1$ MHz.

When a single qubit is on-resonance with the cavity, so that the detuning is $\Delta = \omega - \omega_r = 0$, the individual eigenstates of the cavity ($|0\rangle, |1\rangle$) and the qubit ($|g\rangle, |e\rangle$) are no longer the eigenstates of the coupled system. Here, we find new eigenstates formed by an equal combination of cavity photons and qubit excitations, leading to the symmetric and antisymmetric superpositions $(|0\rangle|e\rangle \pm |1\rangle|g\rangle)/\sqrt{2}$. We also find that the energy-level separation of the new eigenstates, $\hbar(\omega \pm g)$, shows the typical vacuum-Rabi mode splitting. In addition, the exchange of photons between the cavity and a single qubit is strongly on-resonance.

In a familiar cavity QED experiment, an atom is excited, $|e\rangle$, and then interacts with an empty cavity, $|0\rangle$. The initial coupled-system state $|0\rangle|e\rangle$ begins to oscillate in time according to $\cos(gt)|0\rangle|e\rangle - i\sin(gt)|1\rangle|g\rangle$, so that the atom excitation $|e\rangle$ is transformed into a cavity photon $|1\rangle$ after a time $t = \pi/2g$ set by the interaction energy $\hbar g$. This continues coherently, with the photon continuously being transferred back and forth between the atom and the cavity in a process known as vacuum-Rabi oscillations. In typical atom-cavity systems^{14,15}, the interaction time is controlled by the atom's velocity through the cavity. In our analogous phase qubit-cavity system, we have the flexibility of using fast (~ 1 ns rise time), roughly rectangular flux bias shift pulses to adjust the interaction time (pulse width) and detuning Δ (pulse amplitude).

As a first demonstration of strongly coupled circuit QED in our system, these two basic vacuum-Rabi behaviours were independently verified for qubit A and qubit B. In Fig. 2a, we show an example of the vacuum-Rabi splitting for qubit B (a similar splitting was obtained for qubit A) using well-established spectroscopic techniques^{24,26}. Vacuum-Rabi oscillations were also obtained for both qubits using an analogous technique borrowed from quantum optics²⁷ and used previously for a superconducting flux qubit coupled to a lumped-element cavity²¹. With qubit B fixed at a given detuning Δ_B , a fast (~ 4 ns) π pulse was applied to qubit B, preparing the initial coupled-system state $|0\rangle|e\rangle_B$. After a short interaction time, the state of the qubit is measured using a fast bias pulse in a manner identical to previous coupled-phase qubit experiments^{11,26}.

In Fig. 2b, we show an example of vacuum-Rabi oscillations for qubit B (similar oscillations were obtained for qubit A) for various detunings Δ_B with a raw contrast of $\sim 20\%$, visible out to 200 ns. We see an increase in the vacuum-Rabi frequency with detuning, roughly proportional to $\sqrt{4g_B^2 + \Delta_B^2}$, with a minimum value on-resonance ($\Delta_B = 0$). An additional energy splitting, near the cavity resonance (seen in Fig. 2a on the lower spectroscopic branch) caused by a two-level system (TLS) defect common to large-area Josephson phase qubits^{24,26}, is responsible for a slight broadening of the spectroscopic splitting and a beating in the oscillations centred at $\Delta_B/g_B \approx 0.5$. Numerical calculations taking into account the size and position of the TLS agree well with the data for $g_B/\pi \approx 86$ MHz, where a small amount of beating is still visible on-resonance (see the inset of

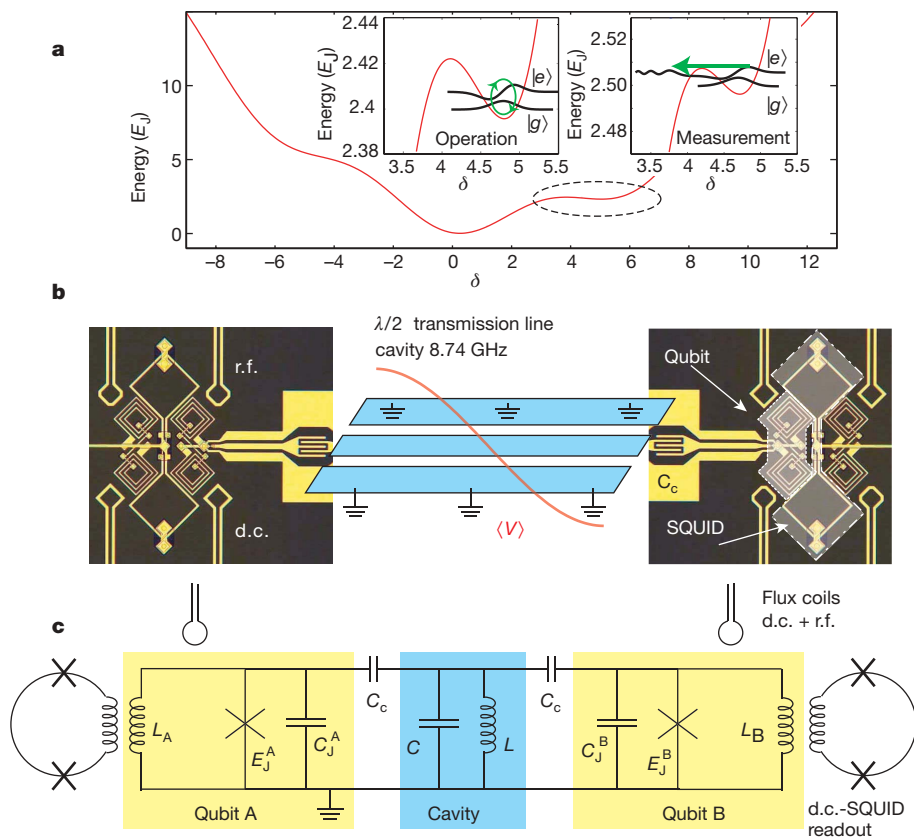


Figure 1 | Schematic description of the experiment set-up. **a**, Potential-energy diagram of the phase qubit (red line) and illustration of the operation and measurement (green line), where tunnelling of the qubit excited state $|e\rangle$ results in a difference of about one flux quantum in the loop, which is read out by a d.c. SQUID²⁴. Repeated simultaneous single-shot measurements¹¹ (typically 1,000 events per data point) provides the necessary statistics to determine the excited-state population P_A and P_B of qubits A and B. **b**, Illustration of the quantum memory element with two Josephson phase qubits (with junction areas $\sim 14 \mu\text{m}^2$) connected via coupling capacitors

$C_c \approx 6.2$ fF to either end of a resonant cavity formed by a 7-mm-long slowly meandering coplanar waveguide, with the qubits separated by about 1.1 mm. The red line depicts the voltage amplitude of the lowest $\lambda/2$ -mode. The device was fabricated with standard optical lithography, producing Al/AIO_x/Al junctions on a sapphire substrate, using SiN_x as an insulator between the metallic layers. r.f., radio frequency. **c**, Lumped element equivalent circuit near the $\lambda/2$ resonance. The cavity has an effective inductance $L \approx 580$ pH and capacitance $C \approx 0.57$ pF, and both qubits had roughly $L_{A,B} \approx 690$ nH, $E_J^{A,B} \approx 45$ K and $C_J^{A,B} \approx 0.7$ pF.

Fig. 2b). Both qubits showed similar behaviour (without a nearby TLS in qubit A), differing by less than 10%, with a coupling strength ($g_A \approx g_B$) matching the design values (see Fig. 1).

After calibrating the amplitude of the shift pulses separately, for both qubits at their far-detuned operation points, we re-measured the vacuum-Rabi oscillations using the shift pulse sequence described in Fig. 2c and confirmed that both experimental methods gave similar results. The raw contrast for the vacuum-Rabi oscillations and for microwave-driven Rabi oscillations on individual qubits ($\sim 40\%$, when far detuned from each other, from the cavity, and from TLS defects) are consistent with previous measurements of coupled interactions using similar phase qubits^{11,26}. It is known that the presence of TLS defects reduces qubit quality and causes contrast loss^{24,26,28}. During flux changes, the qubit frequency sweeps past TLS defects, inducing Landau-Zener transitions to the qubit ground state, reducing measurement contrast.

To investigate the transfer of quantum states through the resonant cavity, we use the vacuum-Rabi interaction of both qubits. The complete sequence (1) to (5) is described in Fig. 2c. Using the static d.c. flux bias coils, the phase qubits are completely detuned ($\Delta_{A,B} \approx 15g_{A,B}$) from the cavity and from each other to suppress any stray cavity and qubit interactions. (1) In this configuration, we prepare a superposition state for qubit A using a rapid microwave pulse. (2) Next, we apply a shift pulse to qubit A, placing it on-resonance with the cavity for a time duration t_A . With shift pulse speeds much greater than $g_A/2\pi$ but still much less than $\omega_A/2\pi$, we effectively preserve the initially prepared quantum state until $\Delta_A = 0$, when the vacuum-Rabi oscillations begin to mix the qubit-cavity states. (3) With the detuning of qubit A restored, we wait for a short storage time $t_S = 10$ ns. (4) Then, we use a second shift pulse to place qubit B

on-resonance with the cavity for a time t_B . (5) Finally, we return qubit B to its fully detuned position and both qubits are measured simultaneously using a fast (~ 4 ns) flux bias measurement pulse^{11,26} that reveals the four joint probabilities $P_{AB} = P_{gg}, P_{eg}, P_{ge}$ and P_{ee} . For simplicity, we choose to focus our attention on the individual excited-state occupation probabilities $P_A = P_{eg} + P_{ee}$ and $P_B = P_{ge} + P_{ee}$.

For the experimental data shown in Fig. 3, we used the state transfer protocol as outlined in Fig. 2c with an initial microwave π pulse applied to qubit A to create a simple pure state $|e\rangle_A$ for transfer. Figure 3a, b shows data over a range of interaction times t_A and t_B . The population maxima (red colour) for P_B in the target qubit B in Fig. 3b satisfy the following conditions: whenever t_A is an odd half-multiple of a vacuum-Rabi period, qubit A has a low population (blue colour) for P_A and we see a corresponding vacuum-Rabi oscillation of P_B occurring in qubit B. The experimental data are in good agreement with theoretical calculations of equation (1) under ideal conditions (Fig. 3c, d).

For clarity, we have extracted a set of three curves from the colour plots of Fig. 3a and b (arrows) and displayed them in Fig. 3e and f. If both shift pulses last for a half-vacuum-Rabi period $\pi/2g_{A,B}$, then the qubit excitation is completely transferred into a cavity photon and the subsequent excited-state population P_A is low, while in the target qubit B, we simultaneously observe clear vacuum-Rabi oscillations (black curve). The fact that the oscillations start from a minimum indicates the presence of a photon in the cavity at the moment of state transfer to qubit B, as expected. Thus, the excitation must leave qubit A, enter the cavity, where it is stored for a short time, and then finally be deposited in qubit B. Repeating this experiment for a full vacuum-Rabi period ($t_A = \pi/g_A \approx 11.6$ ns, green curves) shows no oscillations

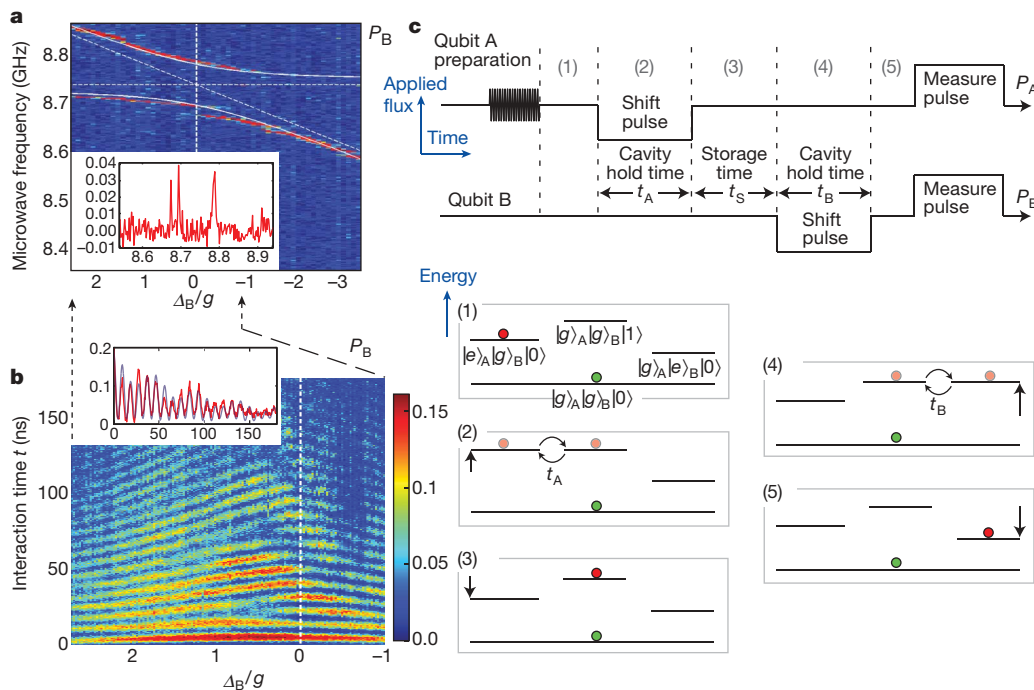


Figure 2 | Demonstration of strongly coupled circuit QED. **a**, Microwave spectroscopy of qubit B showing the vacuum-Rabi splitting (qubit A detuned). Blue colour represents low P_B , red represents high P_B . The inset shows a cross-section at $\Delta_B = 0$ (along the dashed vertical line). **b**, Vacuum-Rabi oscillations in qubit B after a short π pulse. The inset shows a cross-section at $\Delta_B = 0$. The blue line shows the numerical results including the TLS defect at near-resonance. **c**, Illustration of the general quantum state transfer protocol performed by a sequence of flux bias pulses applied to qubits A and B. Here each qubit is effectively decoupled from the cavity, except during the shift pulses, which bring them into resonance with the cavity, one qubit at a time. (1) An arbitrary superposition state $\alpha|g\rangle_A + \beta|e\rangle_A$

is prepared in qubit A. The red and green circles represent mixtures of the occupied energy levels. (2) Qubit A is shifted into resonance with the cavity for an interaction time lasting half of a vacuum-Rabi period, $t_A = \pi/2g_A$, the photon has been exchanged and the state of qubit A has been mapped into a superposition $\alpha|0\rangle + \beta|1\rangle$ of the two lowest-photon-number eigenstates (Fock states) of the cavity. (3) Qubit A is shifted off-resonance, storing the initial state in the cavity for a time duration t_S . (4) Qubit B is shifted into resonance for half of a vacuum-Rabi period, $t_B = \pi/2g_B$, transferring the state into qubit B, leaving the cavity in its ground state $|0\rangle$. (5) Both qubits are detuned, completing the coherent quantum state transfer from qubit A to qubit B.

in P_B , also as expected, because the photon was fully returned to qubit A (as indicated by higher values of P_A), leaving the cavity empty. The red lines illustrate an intermediate case, with $t_A \approx 3/4$ of a vacuum-Rabi period yielding oscillations of lower amplitude but the same frequency. Thus, we conclude that we can clearly transfer photons between two phase qubits, through the resonant cavity, as well as store this quantum information for a short time. By performing single-photon transfers for longer and longer storage times, we find a energy decay time of $\sim 1 \mu\text{s}$ for the cavity. Superconducting cavities are simple and tend to be more coherent than state-of-the-art superconducting qubits. With high-quality-factor superconducting microwave resonators^{14,29}, photon lifetimes near $100 \mu\text{s}$ could provide us with a feasible short-term memory element for use in superconducting quantum information systems.

To verify that quantum coherence is maintained during state transfer for an arbitrary superposition state, we perform a Ramsey fringe-type interference experiment⁷ that preserves the quantum state up to a relative phase factor. We follow a protocol (Fig. 4a) similar to that used previously, except that here, we first prepare qubit A in an equal-weight superposition state $(|g\rangle_A + |e\rangle_A)/\sqrt{2}$, using a $\pi/2$ pulse applied slightly off-resonance, $\Delta\omega_A = \omega_d - \omega_A$, where ω_d is the microwave drive frequency. Again, we perform shift pulses, first, to map the initial state onto a superposition of the two lowest-photon-number states $|0\rangle$ and $|1\rangle$ of the cavity and, second, to retrieve this quantum information through the transfer to the states $|g\rangle_B$ and $|e\rangle_B$ spanned by qubit B. Owing to the short, but finite, time duration of the shift pulses and a brief storage time delay in the cavity, the state transferred to qubit B becomes $(|g\rangle_B + \exp(i\Theta)|e\rangle_B)/\sqrt{2}$ where Θ represents a trivial phase accumulation during the transfer process correctable by single-qubit rotations. Following the coherent state transfer to qubit B, we expect a clear precession of the transferred state during the time delay Δt . By applying a final $\pi/2$ pulse to qubit B (also slightly off-resonance, $\Delta\omega_B = \Delta\omega_A$), we complete the Ramsey fringe-type experiment, rotating qubit B into a different state depending on the total relative phase shift accumulated over the time

Δt . In Fig. 4b and c, we show the expected Ramsey-type oscillations with frequencies linearly proportional to the microwave detuning $\Delta\omega_B$, thus verifying the transfer of quantum coherence through the cavity qubus.

To test the integrity of our experimental design, we investigated in detail the possible role of stray unintended photon generation in the cavity, both d.c. and radio-frequency inductive flux cross-coupling between the two qubits, the role of nearby TLS defects, and measurement cross-talk¹¹ directly through the cavity. First, we verified that the experiment satisfied basic consistency checks based on predictions of the model hamiltonian, equation (1), by altering the transfer pulse sequence shown in Fig. 2c. When we applied π pulses to either qubit with any of the shift pulses omitted, we saw no visible oscillations (above 1% contrast) in the target qubit. When compared to the $\sim 20\%$ contrast of the full state transfer sequence, this corresponds to less than 0.05 stray photons in the cavity per π pulse, while the expected thermal occupation of the cavity with $\hbar\omega_r \approx 420 \text{ mK}$ at $T < 100 \text{ mK}$ should be $n < 0.02$. Next, we determined the cross-coupling of shift pulses by studying the flux modulation of one qubit for flux applied to the other qubit. We found at most a 6% cross-coupling between the two qubits, allowing us to avoid bias pulse cross-talk for large detunings. We performed numerical simulations that included the finite coherence times of each qubit, nearby TLS defects and no additional cross-talk. These results agree with the data, as shown in Fig. 2b. Finally, we performed detailed time-delay measurements¹¹ to investigate the role of measurement cross-talk when qubits A and B were not measured simultaneously. These results show that the cavity acts like an extremely narrow bandpass filter (centred at ω_r) that helps to block either qubit from the broadband transient microwave radiation generated by the measurement process¹¹. Moreover, the use of shift pulses ensures both qubits are far detuned during the measurement process, minimizing excitation errors from radiation at frequency ω_r passing through the cavity.

We have successfully coupled two superconducting Josephson phase qubits through a resonant microwave cavity and have

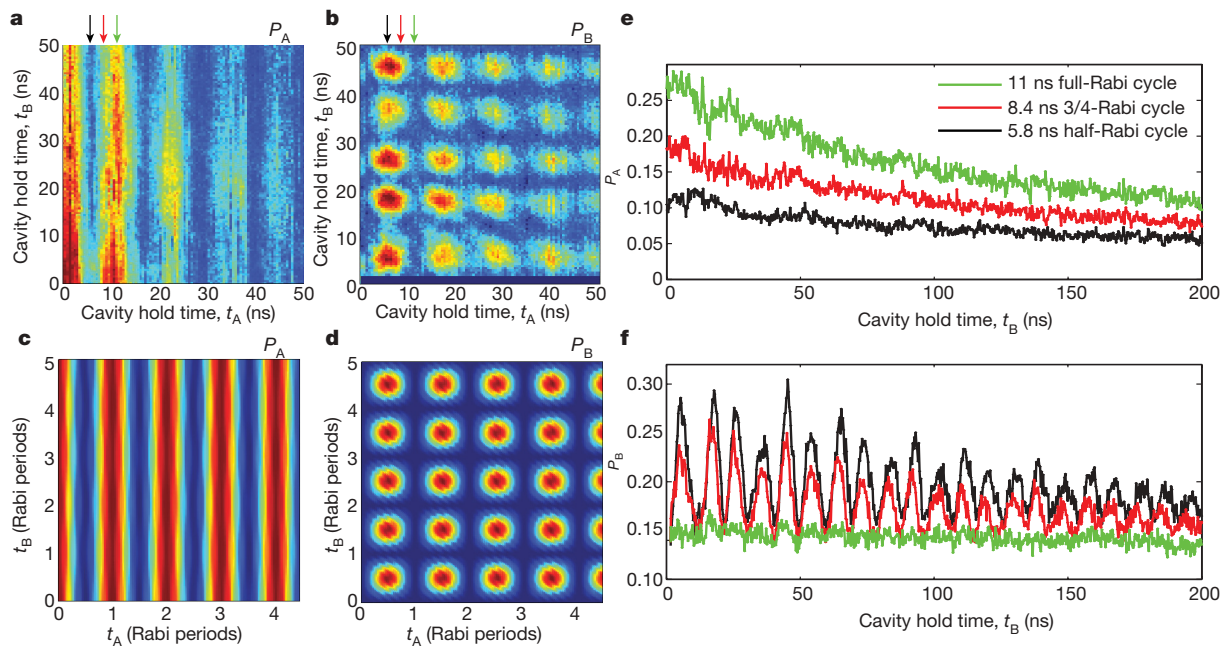


Figure 3 | Experimental data showing the quantum state transfer from qubit A to qubit B via the cavity. The transfer follows the protocol in Fig. 2, where the qubit A excited state $|e\rangle_A$ is first mapped into the single photon state $|1\rangle$ in the cavity, and then transferred into qubit B. **a, b**, Measured populations of qubits A and B as functions of the cavity hold times. Blue colour represents low $P_{A,B}$, red represents high $P_{A,B}$. **c, d**, Corresponding theoretical prediction for ideal conditions without decoherence and 100% fidelity. There is clear agreement between the experimental data and the

ideal predictions, although qubit A was of lower quality than qubit B, most probably owing to fabrication imperfections. **e**, Excited-state occupancy P_A of the source qubit A reveals a lower population if the interaction time equals half of a vacuum-Rabi period, $t_A = \pi/2g_A \approx 5.8 \text{ ns}$ (black line). **f**, Simultaneous measurement of qubit B shows vacuum-Rabi oscillations induced by the transfer of a single photon (black line). Here the black, red and green curves in **e** (or **f**) correspond to data indicated by the arrows in **a** (or **b**); for a full discussion see the text.

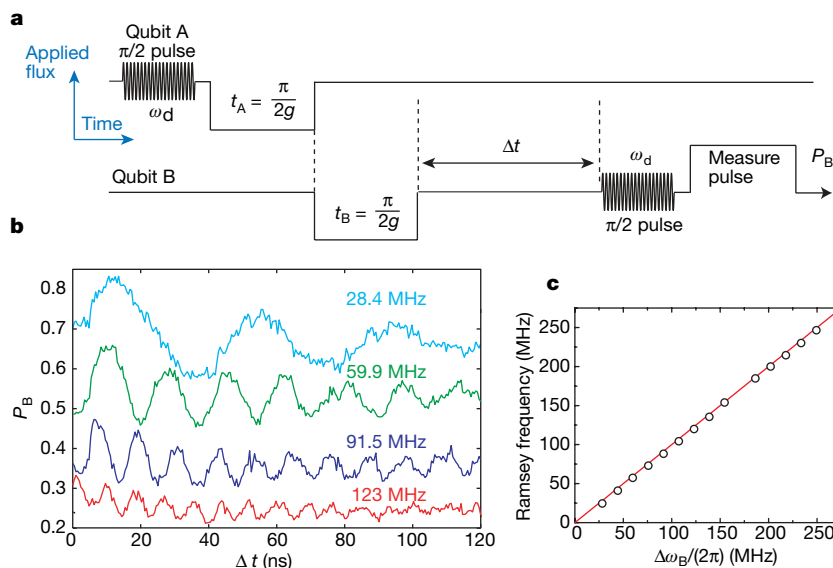


Figure 4 | Demonstration of the coherent transfer of a quantum state through the quantum bus. We use $\pi/2$ microwave pulses from the level spacing frequencies $\omega_{A,B}$ of qubits A and B in order to perform a Ramsey fringe-type interference experiment. **a**, We prepare an equal-weight superposition state $(|g\rangle_A + |e\rangle_A)/\sqrt{2}$ in qubit A using a 10-ns-long $\pi/2$ pulse (with frequency ω_d) while both qubits are detuned from the cavity and from each other. We transfer this state into qubit B as in Fig. 2, and then wait for a delay time Δt before applying a detuned $\pi/2$ pulse to qubit B. This is

observed vacuum-Rabi splittings, vacuum-Rabi oscillations, and the coherent transfer and storage of quantum states mediated by the cavity. The fidelity of the state transfer protocol is mostly limited by the quality of the phase qubits, the presence of TLS defects, and the non-optimization of the shape of the shift pulses performing the state transfer. We intend to improve qubit quality by eliminating dielectric materials and reducing the size of Josephson junctions to remove TLS defects^{28,30}. Further measurements involving full state tomography^{12,30} need to be performed to fully quantify the fidelity of this cavity qubus. This demonstration clearly shows progress towards the storage and communication of quantum information using coherent superconducting systems of multiple qubits, an exciting new frontier for solid-state circuit QED and quantum information science.

Received 18 April; accepted 25 July 2007.

1. Julsgaard, B., Sherson, J., Cirac, J. I., Fiurek, J. & Polzik, E. S. Experimental demonstration of quantum memory for light. *Nature* **432**, 482–486 (2004).
2. Langer, C. *et al.* Long-lived qubit memory using atomic ions. *Phys. Rev. Lett.* **95**, 060502 (2005).
3. Plastina, F. & Falci, G. Communicating Josephson qubits. *Phys. Rev. B* **67**, 224514 (2003).
4. Cleland, A. N. & Geller, M. R. Superconducting qubit storage and entanglement with nano-mechanical resonators. *Phys. Rev. Lett.* **93**, 070501 (2004).
5. Bouchiat, V. *et al.* Quantum coherence with a single Cooper pair. *Phys. Scr.* **T76**, 165–170 (1998).
6. Nakamura, Y., Pashkin, Yu. A. & Tsai, J. S. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature* **398**, 786–788 (1999).
7. Vion, D. *et al.* Manipulating the quantum state of an electrical circuit. *Science* **296**, 886–889 (2002).
8. Chiorescu, I., Nakamura, Y., Harmans, C. J. & Mooij, J. E. Coherent quantum dynamics of a superconducting flux qubit. *Science* **299**, 1869–1871 (2003).
9. Makhlin, Y., Schön, G. & Shnirman, A. Quantum-state engineering with Josephson-junction devices. *Rev. Mod. Phys.* **73**, 357–400 (2001).
10. Yamamoto, T., Pashkin, Y. A., Astafiev, O., Nakamura, Y. & Tsai, J. S. Quantum oscillations in two coupled charge qubits. *Nature* **425**, 941–944 (2003).
11. McDermott, R. *et al.* Simultaneous state measurement of coupled Josephson phase qubits. *Science* **307**, 1299–1302 (2005).
12. Steffen, M. *et al.* Measurement of the entanglement of two superconducting qubits via state tomography. *Science* **313**, 1423–1425 (2006).
13. Grajcar, M. *et al.* Four-qubit device with mixed couplings. *Phys. Rev. Lett.* **96**, 047006 (2006).

analogous to Ramsey fringe experiments with single qubits, where a coherent quantum state slowly precesses at the microwave detuning frequency $\Delta\omega_B = \omega_B - \omega_d$. **b**, Coherent oscillations in qubit B for several detunings (vertically displaced for clarity). **c**, The frequency of the Ramsey-type oscillations as a function of the microwave detuning. The solid red line represents the theoretical predictions with no fitting parameters; circles are experimental data points.

14. Schleich, W. P. & Walther, H. *Elements of Quantum Information* 1st edn, Ch. 1 (Wiley-VCH, New York, 2007).
15. Haroche, S. & Raimond, J.-M. *Exploring the Quantum: Atoms, Cavities, and Photons* 1st edn, Ch. 5 (Oxford Univ. Press, Oxford, 2006).
16. Maître, X. *et al.* Quantum memory with a single photon in a cavity. *Phys. Rev. Lett.* **79**, 769–772 (1997).
17. Buisson, O. & Hekking, F. W. J. in *Macroscopic Quantum Coherence and Quantum Computing* (eds Averin, D. V., Ruggiero, B. & Silvestrini, P.) 137–145 (Kluwer Academic, New York, 2001).
18. Blais, A., Huang, R.-S., Wallraff, A., Girvin, S. M. & Schoelkopf, R. J. Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation. *Phys. Rev. A* **69**, 062320 (2004).
19. Wallraff, A. *et al.* Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature* **431**, 162–167 (2004).
20. Chiorescu, I. *et al.* Coherent dynamics of a flux qubit coupled to a harmonic oscillator. *Nature* **431**, 159–162 (2004).
21. Johansson, J. *et al.* Vacuum Rabi oscillations in a macroscopic superconducting qubit LC oscillator system. *Phys. Rev. Lett.* **96**, 127006 (2006).
22. Koch, R. H. *et al.* Experimental demonstration of an oscillator stabilized Josephson flux qubit. *Phys. Rev. Lett.* **96**, 127001 (2006).
23. Xu, H. *et al.* Spectroscopy of three-particle entanglement in a macroscopic superconducting circuit. *Phys. Rev. Lett.* **94**, 027003 (2005).
24. Simmonds, R. W. *et al.* Decoherence in Josephson phase qubits from junction resonators. *Phys. Rev. Lett.* **93**, 077003 (2004).
25. Martinis, J. M., Nam, S., Aumentado, J. & Urbina, C. Rabi oscillations in a large Josephson-junction qubit. *Phys. Rev. Lett.* **89**, 117901 (2002).
26. Cooper, K. B. *et al.* Observation of quantum oscillations between a Josephson phase qubit and a microscopical resonator using fast readout. *Phys. Rev. Lett.* **93**, 180401 (2004).
27. Brune, M. *et al.* Quantum Rabi oscillation: A direct test of field quantization in a cavity. *Phys. Rev. Lett.* **76**, 1800–1803 (1996).
28. Martinis, J. M. *et al.* Decoherence in Josephson qubits from dielectric loss. *Phys. Rev. Lett.* **95**, 210503 (2005).
29. Day, P. K., LeDuc, H. G., Mazin, B. A., Vayonakis, A. & Zmuidzinas, J. A broadband superconducting detector suitable for use in large arrays. *Nature* **425**, 817–821 (2003).
30. Steffen, M. *et al.* State tomography of capacitively shunted phase qubits with high fidelity. *Phys. Rev. Lett.* **97**, 050502 (2006).

Acknowledgements We gratefully acknowledge discussions with J. Aumentado, K. Cicak, K. Osborne, R. Schoelkopf and D. Wineland. This work was financially supported by the NIST and the DTO.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.W.S. (simmonds@boulder.nist.gov).

Coupling superconducting qubits via a cavity bus

J. Majer^{1*}, J. M. Chow^{1*}, J. M. Gambetta¹, Jens Koch¹, B. R. Johnson¹, J. A. Schreier¹, L. Frunzio¹, D. I. Schuster¹, A. A. Houck¹, A. Wallraff^{1†}, A. Blais^{1†}, M. H. Devoret¹, S. M. Girvin¹ & R. J. Schoelkopf¹

Superconducting circuits are promising candidates for constructing quantum bits (qubits) in a quantum computer; single-qubit operations are now routine^{1,2}, and several examples^{3–9} of two-qubit interactions and gates have been demonstrated. These experiments show that two nearby qubits can be readily coupled with local interactions. Performing gate operations between an arbitrary pair of distant qubits is highly desirable for any quantum computer architecture, but has not yet been demonstrated. An efficient way to achieve this goal is to couple the qubits to a ‘quantum bus’, which distributes quantum information among the qubits. Here we show the implementation of such a quantum bus, using microwave photons confined in a transmission line cavity, to couple two superconducting qubits on opposite sides of a chip. The interaction is mediated by the exchange of virtual rather than real photons, avoiding cavity-induced loss. Using fast control of the qubits to switch the coupling effectively on and off, we demonstrate coherent transfer of quantum states between the qubits. The cavity is also used to perform multiplexed control and measurement of the qubit states. This approach can be expanded to more than two qubits, and is an attractive architecture for quantum information processing on a chip.

There are several physical systems in which one could realize a quantum bus. A particular example is trapped ions^{10,11} in which a variety of quantum operations and algorithms have been performed using the quantized motion of the ions (phonons) as the bus. Photons are another natural candidate as a carrier of quantum information^{12,13}, because they are highly coherent and can mediate interactions between distant objects. To create a photon bus, it is helpful to use the increased interaction strength provided by the techniques of cavity quantum electrodynamics (QED), where an atom is coupled to a single-cavity mode. In the strong coupling limit¹⁴ the interaction is coherent, permitting the transfer of quantum information between the atom and the photon. Entanglement between atoms has been demonstrated with Rydberg atom cavity QED^{15–17}. Circuit QED¹⁸ is a realization of the physics of cavity QED with superconducting qubits coupled to a microwave cavity on a chip. Previous circuit QED experiments with single qubits have achieved¹⁹ the strong coupling limit and have demonstrated²⁰ the transfer of quantum information from qubit to photon. Here we perform a circuit QED experiment with two qubits strongly coupled to a cavity, and demonstrate a coherent, non-local coupling between the qubits via this bus.

Operations with multiple superconducting qubits have been performed and are a subject of current research. The first solid-state quantum gate has been demonstrated with charge qubits³. For flux qubits, two-qubit coupling⁵ and a controllable coupling mechanism have been realized^{7–9}. Two phase qubits have also been successfully coupled⁴ and the entanglement between them has been observed⁶. All of these interactions have been realized by connecting qubits via lumped circuit elements, with capacitive coupling in the case of

charge and phase qubits, and inductive coupling for flux qubits. Therefore, these coupling mechanisms have been restricted to local interactions and couple only nearest-neighbour qubits. In this work, we present a coupling that is realized with a cavity that is a distributed circuit element, rather than with the lumped elements used previously. The interaction between the qubits occurs via photons in the cavity; hence, the cavity acts as an interaction bus allowing a non-local coupling of multiple qubits.

To realize the cavity bus, we place two superconducting qubits 5 mm apart at opposite ends of a superconducting transmission line resonator (Fig. 1a, b). The qubits are transmons²¹, a charge-phase qubit that is a modified version of the Cooper pair box. In this type of qubit, the Josephson energy is larger than the charging energy ($E_J \gg E_C$) and the transition frequency between the ground state and the first excited state is given by $\omega \approx \sqrt{8E_J E_C}/\hbar$. The Josephson junctions are arranged in a split-pair geometry, so that the Josephson energy, $E_J = E_J^{\max} |\cos(\pi\Phi/\Phi_0)|$ depends on the magnetic flux Φ applied through the split-pair loop. Hence, the transition frequency of the qubits, $\omega_{1,2} = \omega_{1,2}^{\max} \sqrt{|\cos(\pi\Phi/\Phi_0)|}$ can be tuned *in situ* with the applied flux. The size of the two loops is different and incommensurate, so that control of the two transition frequencies is attainable with a certain degree of independence. To probe the state of the system, homodyne detection of the transmitted signal is performed and both quadrature voltages are recorded, which allows reconstruction of the phase and amplitude of the transmitted signal.

In the first measurement we observe strong coupling of each of the qubits separately to the cavity. By varying the flux, each of the two qubits can be tuned into resonance with the cavity (see Fig. 2a). Whenever a qubit and the cavity are degenerate, the transmission is split into two well-resolved peaks in frequency, an effect called vacuum Rabi splitting¹⁹, demonstrating that each qubit is in the strong coupling limit with the cavity. Each of the peaks corresponds to a superposition of qubit excitation and a cavity photon in which the energy is shared between the two systems. From the frequency difference at the maximal splitting, the coupling parameters $g_{1,2} \approx 105$ MHz can be determined for each qubit. The transition frequency of each of the two qubits (see Fig. 2a) can also be measured far from the cavity frequency, as described below.

In the remainder of the experiment we operate the system in the dispersive limit, where both qubits are detuned from the resonator ($|\mathcal{A}_{1,2}| = |\omega_{1,2} - \omega_r| \gg g_{1,2}$). In this limit, we use second-order perturbation theory, and the full system with the two qubits and the cavity is described by the effective hamiltonian²²:

$$H_{\text{eff}} = \frac{\hbar\omega_1}{2} \sigma_1^z + \frac{\hbar\omega_2}{2} \sigma_2^z + \hbar(\omega_r + \chi_1 \sigma_1^z + \chi_2 \sigma_2^z) a^\dagger a + \hbar J (\sigma_1^- \sigma_2^+ + \sigma_2^- \sigma_1^+) \quad (1)$$

In this regime, no energy is exchanged with the cavity. However, the qubits and cavity are still dispersively coupled, resulting in a qubit-state-dependent shift $\pm\chi_{1,2}$ of the cavity frequency (see

¹Departments of Applied Physics and Physics, Yale University, New Haven, Connecticut 06520, USA. [†]Present addresses: Department of Physics, ETH Zurich, CH-8093 Zürich, Switzerland (A.W.); Département de Physique et Regroupement Québécois sur les Matériaux de Pointe, Université de Sherbrooke, Sherbrooke, Québec, J1K2R1 Canada (A.B.).

*These authors contributed equally to this work.

Fig. 3a) or equivalently an alternating current (a.c.) Stark shift of the qubit frequencies²³. The frequency shift $\chi_{1,2}$ can be calculated from the detuning $\Delta_{1,2}$ and the measured coupling strength $g_{1,2}$ (ref. 21). The last term describes the interaction between the qubits, which is a transverse exchange interaction of strength $J = g_1 g_2 (1/\Delta_1 + 1/\Delta_2)/2$ (see Fig. 1c). The qubit–qubit interaction is a result of virtual exchange of photons with the cavity. When the qubits are degenerate with each other, an excitation in one qubit can be transferred to the other qubit by virtually becoming a photon in the cavity (see Fig. 3b). However, when the qubits are non-degenerate $|\omega_1 - \omega_2| \gg J$, this process does not conserve energy, and therefore the interaction is effectively turned off. Thus, instead of modifying the actual coupling constant^{7–9}, we control the effective coupling strength by tuning the qubit transition frequencies. This is possible because the qubit–qubit coupling is transverse, which also distinguishes our experiment from the situation in liquid-state nuclear magnetic resonance (NMR) quantum computation, where an effective switching-off can only be achieved by repeatedly applying decoupling pulses²⁴.

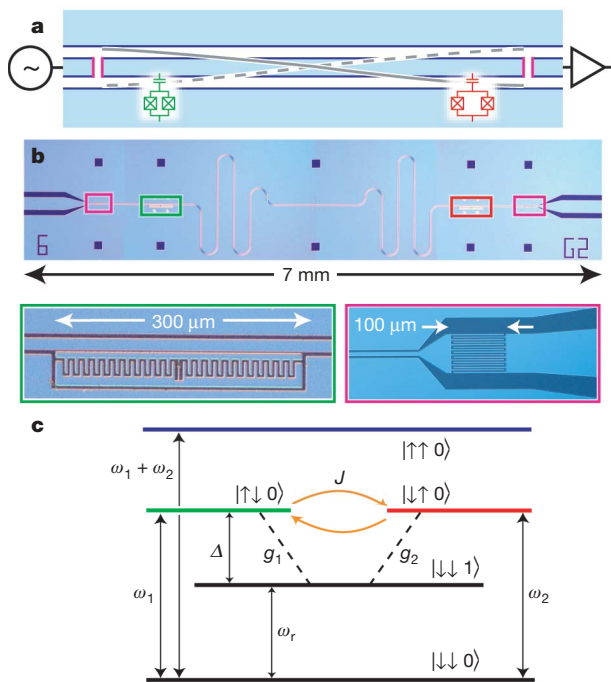


Figure 1 | Sample and scheme used to couple two qubits to an on-chip microwave cavity. Circuit (a) and optical micrograph (b) of the chip with two transmon qubits coupled by a microwave cavity. The cavity is formed by a coplanar waveguide (light blue) interrupted by two coupling capacitors (purple boxes and inset). The resonant frequency of the cavity is $\omega_r/2\pi = 5.19$ GHz and its width is $\kappa/2\pi = 33$ MHz, determined by the coupling capacitors. The cavity is operated as a half-wave resonator ($L = \lambda/2 = 12.3$ mm) and the electric field in the cavity is indicated by the grey line. The two transmon qubits (optimized Cooper pair boxes, red and green boxes and inset) are located at opposite ends of the cavity, where the electric field has an antinode. Each transmon qubit consists of two superconducting islands connected by a pair of Josephson junctions and an extra shunting capacitor ('interdigitated finger' structure in the green inset). The left qubit (qubit 1) has a charging energy of $E_{C1}/h = 424$ MHz and maximum Josephson energy of $E_{J1}^{\max}/h = 14.9$ GHz. The right qubit (qubit 2) has a charging energy $E_{C2}/h = 442$ MHz and maximum Josephson energy of $E_{J2}^{\max}/h = 18.9$ GHz. The loop area between the Josephson junctions for the two transmon qubits differs by a factor of approximately 5/8, allowing a differential flux bias. The microwave signals enter the chip from the left, and the response of the cavity is amplified and measured on the right. **c**, Scheme of the dispersive qubit–qubit coupling. When the qubits are detuned from the cavity ($|\Delta_{1,2}| = |\omega_{1,2} - \omega_r| \gg g_{1,2}$) the qubits both dispersively shift the cavity. The excited state in the left qubit $|\uparrow\downarrow 0\rangle$ interacts with the excited state in the right qubit $|\uparrow\uparrow 0\rangle$ via the exchange of a virtual photon $|\downarrow\downarrow 1\rangle$ in the cavity.

We first observe the coherent interaction between the two qubits via the cavity by performing spectroscopy of their transition frequencies (see Fig. 2). This is done by monitoring the change in cavity transmission when the qubits are probed by a second microwave signal. By applying a magnetic flux the qubits can be tuned through resonance with each other (see Fig. 2b), revealing an avoided crossing. The magnitude of the splitting agrees well with the theoretical value $2J = 2g_1 g_2 / \Delta = 2\pi \cdot 26$ MHz when one takes into account that $g_{1,2}$ vary with frequency for a transmon qubit²¹. The splitting is well resolved, with a magnitude J much greater than the qubit linewidths, indicating a coherent coupling and that the qubits are in the strong dispersive limit²⁵. We note that although the coupling strength J is smaller than the cavity decay rate $\kappa/2\pi \approx 33$ MHz, the avoided crossing is nearly unaffected by the cavity loss. This is possible in such a large- κ cavity, which is required for fast measurements, because only virtual photons are exchanged; if real photons were used, the cavity-induced relaxation of the qubits (Purcell effect²⁰) would make coherent state transfer unfeasible.

Another manifestation of the coherence of this interaction is the observation of a dark state. One observes a disappearance of the spectroscopic signal near the crossing point, which is due to

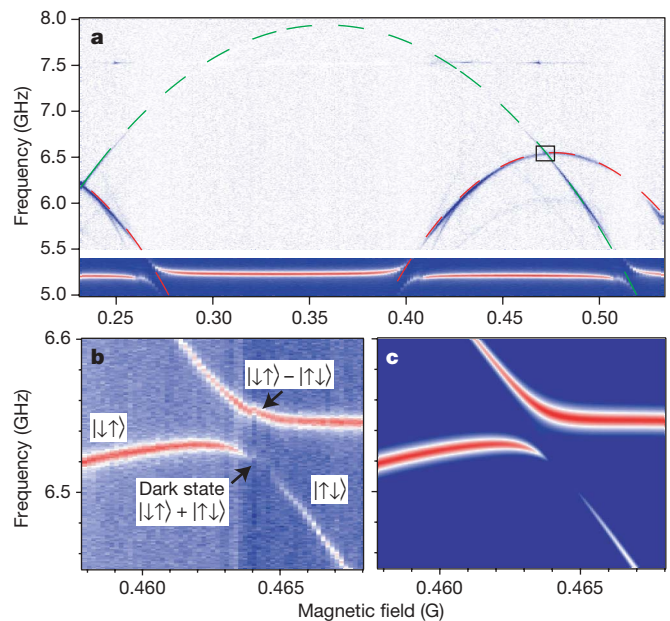


Figure 2 | Cavity transmission and spectroscopy of single and coupled qubits. **a**, The transmission through the cavity as a function of applied magnetic field is shown in the frequency range between 5 and 5.4 GHz. When either of the qubits is in resonance with the cavity, the cavity transmission shows an avoided crossing due to the vacuum Rabi splitting. The maximal vacuum Rabi splitting for the two qubits is the same within the measurement uncertainty and is ~ 105 MHz. Above 5.5 GHz, spectroscopic measurements of the two qubit transitions are displayed. A second microwave signal is used to excite the qubit and the dispersive shift of the cavity frequency is measured. The dashed lines show the resonance frequencies of the two qubits, which are a function of the applied flux according to $\omega_{1,2} = \omega_{1,2}^{\max} \sqrt{|\cos(\pi\Phi/\Phi_0)|}$. The maximum transition frequency for the first qubit is $\omega_1^{\max}/2\pi = 7.8$ GHz and for the second qubit is $\omega_2^{\max}/2\pi = 6.45$ GHz. For strong drive powers, additional resonances between higher qubit levels are visible. The black box indicates the region shown in **b**. **b**, Spectroscopy of the two-qubit crossing. The qubit levels show a clear avoided crossing with a minimum distance of $2J/2\pi = 26$ MHz. At the crossing the eigenstates of the system are symmetric and antisymmetric superpositions of the two qubit states. The spectroscopic drive is antisymmetric and therefore unable to drive any transitions to the symmetric state, resulting in a dark state. **c**, Predicted spectroscopy at the qubit–qubit crossing using a markovian master equation that takes into account higher modes of the cavity. The parameters for this calculation are obtained from the vacuum Rabi splitting and the single-qubit spectroscopy.

destructive interference associated with the fact that the qubits are separated by half a wavelength. At the crossing, the eigenstates are superpositions of the single-qubit states. In particular, the state with lower frequency is the symmetric triplet state $|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle$ and the state at higher frequency is the antisymmetric singlet state $|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle$. In

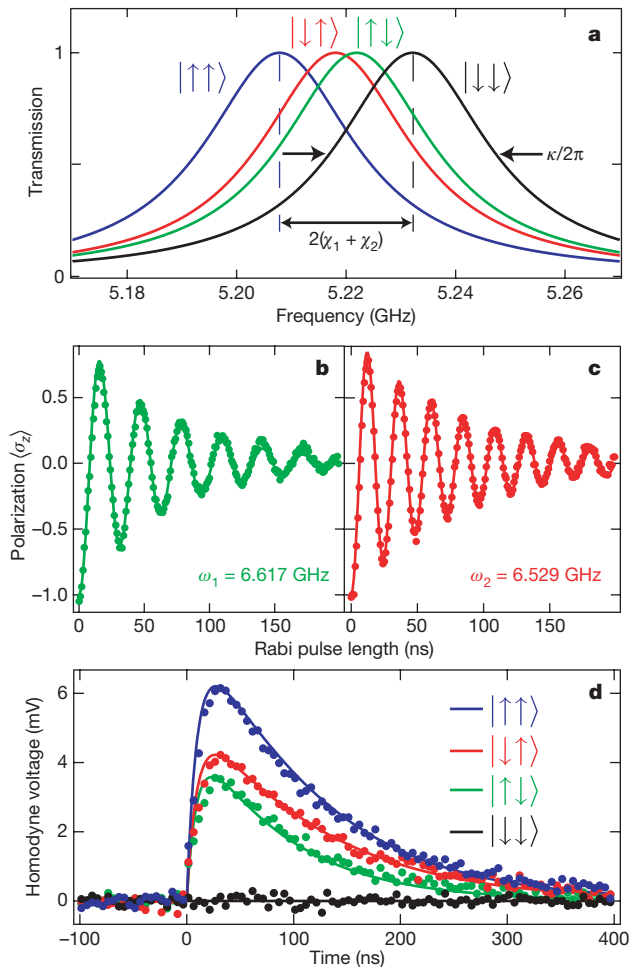


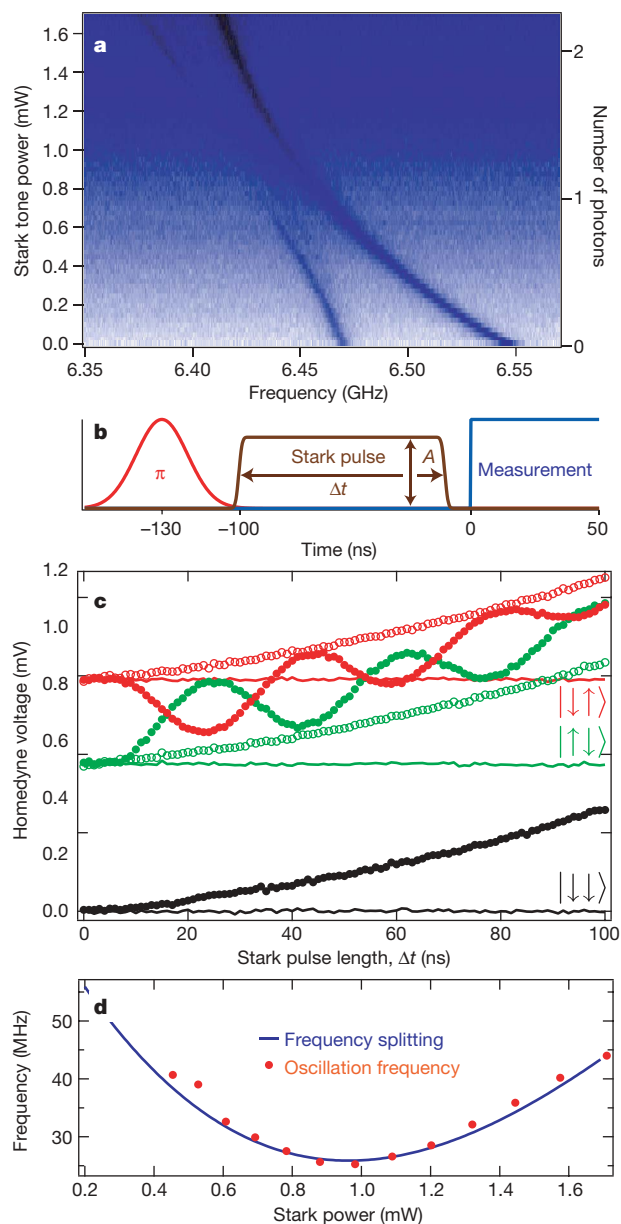
Figure 3 | Multiplexed control and read-out of uncoupled qubits. **a**, Predicted cavity transmission for the four uncoupled qubit states. In the dispersive limit ($|A_{1,2}| = |\omega_{1,2} - \omega_r| \gg g_{1,2}$) the frequency is shifted by $\chi_1\sigma_1^z + \chi_2\sigma_2^z$. Operating the qubits at transition frequencies $\omega_1/2\pi = 6.617$ GHz and $\omega_2/2\pi = 6.529$ GHz, we find $\chi_1/2\pi = -5.9$ MHz and $\chi_2/2\pi = -7.4$ MHz. Measurement is achieved by placing a probe at a frequency where the four cavity transmissions are distinguishable. The two-qubit state can then be reconstructed from the homodyne measurement of the cavity. Rabi oscillations of qubit 1 (**b**) and qubit 2 (**c**) are shown. A drive pulse of increasing duration is applied at the qubit transition frequency and the response of the cavity transmission is measured after the pulse is turned off. Oscillations of quadrature voltages are measured for each of the qubits and mapped onto the polarization $\langle\sigma_z\rangle$. The solid line shows results from a master equation simulation, which takes into account the full dynamics of the two qubits and the cavity. The absence of beating in both traces is a signature of the suppression of the qubit–qubit coupling at this detuning. **d**, The homodyne response (average of 1,000,000 traces) of the cavity after a π -pulse on qubit 1 (green), qubit 2 (red), and both qubits (blue). The black trace shows the level when no pulses are applied. The contrasts²⁸ (that is, the amplitude of the pulse relative to its ideal maximum value) for these pulses are 60% (green), 61% (red) and 65% (blue). The solid line shows the simulated value, including the qubit relaxation and the turn-on time of the cavity. The agreement between the theoretical prediction and the data indicates that the measured contrast is the maximum observable. From the theoretical calculation we can estimate the selectivity (see text for details) for each π -pulse to be 87% (qubit 1) and 94% (qubit 2). We note that this figure of merit is not at all intrinsic and that it could be improved by increasing the detuning between the two qubits for instance, or using shaped excitation pulses.

the dispersive limit, the spectroscopic excitation is of the form $\sigma_1^x g_1/A_1 - \sigma_2^x g_2/A_2$, where the negative sign is due to the opposite signs of the electric field at the different ends of the $\lambda/2$ cavity, as shown in Fig. 1a and b. Thus, such an external signal applied to the cavity cannot drive any transitions to the symmetric state, and is therefore dark. Moreover, just as the triplet state does not couple to the drive, it is protected against decay through the cavity. Conversely, the decay from the singlet state is enhanced, similar to super-radiant effects observed in atomic physics^{26,27}. Figure 2c shows the calculated spectroscopy at the qubit–qubit crossing, which reproduces all qualitative features of the measured data.

In addition to acting as a quantum bus, the cavity can in principle also be used for multiplexed read-out and control of the two qubits. Here, ‘multiplexed’ refers to acquisition of information or control of more than one qubit via a single channel. To address the qubits independently, the flux is tuned such that the qubit frequencies are 88 MHz apart ($\omega_1 = 6.617$ GHz, $\omega_2 = 6.529$ GHz), making the qubit–qubit coupling negligible. Rabi experiments showing individual control are performed by applying a radio-frequency pulse at the resonant frequency of either qubit, followed by a measurement pulse at the resonator frequency. The response (see Fig. 3b and c) is consistent with that of a single-qubit oscillation and shows no beating, indicating that the coupling does not affect single-qubit operations and read-out. With similar measurements the relaxation times T_1 of the two qubits are determined to be 78 and 120 ns, and with Ramsey measurements the coherence times T_2 are found to be 120 and 160 ns. The ability to simultaneously read-out the states of both qubits using a single line is shown by measuring the cavity phase shift, proportional to $\chi_1\sigma_1^z + \chi_2\sigma_2^z$ (see equation (1)), after applying a π -pulse to one or both of the qubits. Figure 3d shows the response of the cavity after a π -pulse has been applied on the first qubit (green points), on the second qubit (red points) or on both qubits (blue points). For comparison the response of the cavity without any pulse applied (black points) is shown. The cavity frequency shifts for the two qubits are different ($\chi_1 \neq \chi_2$), so we are able to distinguish the four states $|\downarrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, and $|\uparrow\uparrow\rangle$ of the qubits with a single read-out line. One can show that this measurement, with a sufficient signal-to-noise ratio and combined with single-qubit rotations, should in principle allow for a full reconstruction of the density matrix (state tomography)⁶, although this is not demonstrated in the present experiment.

The solid lines in Fig. 3d show the results from a theoretical calculation taking into account the full dynamics of the cavity and the two qubits, including the relaxation in the qubits. The agreement of the theory with the measured response shows that the measured contrast is the maximum expected. From the calculated values one can estimate the selectivity, that is, the ability to address one qubit without affecting the other, $S = (P_a - P_u)/(P_a + P_u)$, where P_a and P_u are the maximum populations in the excited state of the addressed qubit and in the excited state of the unaddressed qubit, respectively. The selectivity for qubit 1 is 87% and for qubit 2 is 94%, which indicates good individual control of the qubits.

We can perform coherent state transfer in the time domain by rapidly turning the effective qubit–qubit coupling on and off. Rather than the slow flux tuning discussed above, we now make use of a strongly detuned radio-frequency drive²², which results in an off-resonant Stark shift of the qubit frequencies on the nanosecond timescale. Figure 4a shows the spectroscopy of the two qubits when this off-resonant Stark drive is applied with increasing power. The qubit frequencies are pushed into resonance and a similar avoided crossing is observed, as in Fig. 2b. With the Stark drive’s ability to tune the qubits quickly into resonance, it is possible to observe coherent oscillations between the qubits, using the following protocol (see Fig. 4b): initially the qubits are 80 MHz detuned from each other, where their effective coupling is small, and they are allowed to relax to the ground state $|\downarrow\downarrow\rangle$. Next, a π -pulse is applied to one of the qubits to either create the state $|\downarrow\uparrow\rangle$ or $|\uparrow\downarrow\rangle$. Then a Stark



pulse of power $P_{a.c.}$ is applied bringing the qubits into resonance for a variable time Δt . Because $|\downarrow\uparrow\rangle$ and $|\uparrow\downarrow\rangle$ are not eigenstates of the coupled system, oscillations between these two states occur, as shown in Fig. 4c. Figure 4d shows the frequency of these oscillations for different powers $P_{a.c.}$ of the Stark pulse, which agrees with the frequency domain measurement of the frequency splitting observed in Fig. 4a. These data are strong evidence that the oscillations are due to the coupling between the qubits and that the state of the qubits is transferred from one to the other. A quarter period of these oscillations should correspond to a \sqrt{i} SWAP, which would be a universal gate. Future experiments will seek to demonstrate the performance and accuracy of this state transfer.

The observed qubit–qubit avoided crossing and the coherent state transfer demonstrate that the cavity can act as a coupling bus for superconducting qubits. The interaction is coherent and effectively switchable; furthermore, the coupling is long-range, can easily be extended to non-nearest neighbours and is protected against loss in the bus by the use of virtual photons. This architecture is not restricted to two qubits because there is room to couple many more qubits to the cavity, opening up new possibilities for quantum information processing on a chip.

Figure 4 | Controllable effective coupling and coherent state transfer via off-resonant Stark shift. **a**, Spectroscopy of qubits versus applied Stark tone power. Taking into account an attenuation of 67 dB before the cavity and the filtering effect of the cavity, 0.77 mW corresponds to an average of one photon in the resonator. The qubit transition frequencies (starting at $\omega_1/2\pi = 6.469$ GHz and $\omega_2/2\pi = 6.546$ GHz) are brought into resonance with a Stark pulse applied at 6.675 GHz. An avoided crossing is observed with one of the qubit transition levels becoming dark as in Fig. 2. **b**, Protocol for the coherent state transfer using the Stark shift. The pulse sequence consists of a Gaussian-shaped π -pulse (red) on one of the qubits at its transition frequency $\omega_{1,2}$ followed by a Stark pulse (brown) of varying duration Δt and amplitude A detuned from the qubits, and finally a square measurement pulse (blue) at the cavity frequency. The time between the π -pulse and the measurement is kept fixed at 130 ns. **c**, Coherent state transfer between the qubits according to the protocol above. The plot shows the measured homodyne voltage (average of 3,000,000 traces) with the π -pulse applied to qubit 1 (green dots) and to qubit 2 (red dots) as a function of the Stark pulse length Δt . For reference, the black dots show the signal without any π -pulse applied to either qubit. The overall increase of the signal is caused by the residual Rabi driving due to the off-resonant Stark tone, which is also reproduced by the theory. Improved designs featuring different coupling strengths for the individual qubits could easily avoid this effect. The thin solid lines show the signal in the absence of a Stark pulse. Adding the background trace (black dots) to these, we construct the curves consisting of open circles, which correctly reproduce the upper and lower limits of the oscillating signals owing to coherent state transfer. **d**, The oscillation frequency (red) of the time domain state transfer measurement (**c**) and the splitting frequency (blue) of the continuous-wave spectroscopy (**a**) versus power of the Stark tone. The agreement shows that the oscillations are indeed due to the coupling between the qubits.

Received 24 May; accepted 16 August 2007.

- Devoret, M. H. & Martinis, J. M. Implementing qubits with superconducting integrated circuits. *Quantum Inform. Process.* **3**, 163–203 (2004).
- You, J. Q. & Nori, F. Superconducting circuits and quantum information. *Phys. Today* **58**, 42–47 (2005).
- Yamamoto, T., Pashkin, Y. A., Astafiev, O., Nakamura, Y. & Tsai, J. S. Demonstration of conditional gate operation using superconducting charge qubits. *Nature* **425**, 941–944 (2003).
- Berkley, A. J. *et al.* Entangled macroscopic quantum states in two superconducting qubits. *Science* **300**, 1548–1550 (2003).
- Majer, J. B., Pava, F. G., ter Haar, A. C. J., Harman, C. J. P. M. & Mooij, J. E. Spectroscopy on two coupled superconducting flux qubits. *Phys. Rev. Lett.* **94**, 090501 (2005).
- Steffen, M. *et al.* Measurement of the entanglement of two superconducting qubits via state tomography. *Science* **313**, 1423–1425 (2006).
- Hime, T. *et al.* Solid-state qubits with current-controlled coupling. *Science* **314**, 1427–1429 (2006).
- van der Ploeg, S. H. W. *et al.* Controllable coupling of superconducting flux qubits. *Phys. Rev. Lett.* **98**, 057004 (2007).
- Niskanen, A. O. *et al.* Quantum coherent tunable coupling of superconducting qubits. *Science* **316**, 723–726 (2007).
- Cirac, J. I. & Zoller, P. Quantum computations with cold trapped ions. *Phys. Rev. Lett.* **74**, 4091–4094 (1995).
- Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281–324 (2003).
- Duan, L. M., Lukin, M. D., Cirac, J. I. & Zoller, P. Long-distance quantum communication with atomic ensembles and linear optics. *Nature* **414**, 413–418 (2001).
- Chou, C.-W. *et al.* Functional quantum nodes for entanglement distribution over scalable quantum networks. *Science* **316**, 1316–1320 (2007).
- Mabuchi, H. & Doherty, A. C. Cavity quantum electrodynamics: coherence in context. *Science* **298**, 1372–1377 (2002).
- Hagley, E. *et al.* Generation of Einstein-Podolsky-Rosen pairs of atoms. *Phys. Rev. Lett.* **79**, 1–5 (1997).
- Zheng, S.-B. & Guo, G.-C. Efficient scheme for two-atom entanglement and quantum information processing in cavity QED. *Phys. Rev. Lett.* **85**, 2392–2395 (2000).
- Osnaghi, S. *et al.* Coherent control of an atomic collision in a cavity. *Phys. Rev. Lett.* **87**, 037902 (2001).
- Blais, A., Huang, R.-S., Wallraff, A., Girvin, S. M. & Schoelkopf, R. J. Cavity quantum electrodynamics for superconducting electrical circuits: an architecture for quantum computation. *Phys. Rev. A* **69**, 062320 (2004).
- Wallraff, A. *et al.* Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature* **431**, 162–167 (2004).

20. Houck, A. A. *et al.* Generating single microwave photons in a circuit. *Nature* **449**, 328–331 (2007).
21. Koch, J. *et al.* Charge insensitive qubit design derived from the Cooper Pair Box. *Phys. Rev. A* (in the press); preprint at (<http://arxiv.org/abs/cond-mat/0703002>) (2007).
22. Blais, A. *et al.* Quantum-information processing with circuit quantum electrodynamics. *Phys. Rev. A* **75**, 032329 (2007).
23. Schuster, D. I. *et al.* ac Stark shift and dephasing of a superconducting qubit strongly coupled to a cavity field. *Phys. Rev. Lett.* **94**, 123602 (2005).
24. Gershenfeld, N. A. & Chuang, I. L. Bulk spin-resonance quantum computation. *Science* **275**, 350–356 (1997).
25. Schuster, D. I. *et al.* Resolving photon number states in a superconducting circuit. *Nature* **445**, 515–518 (2007).
26. Grangier, P., Aspect, A. & Vigue, J. Quantum interference effect for two atoms radiating a single photon. *Phys. Rev. Lett.* **54**, 418–421 (1985).
27. Itano, W. M. *et al.* Complementarity and Young's interference fringes from two atoms. *Phys. Rev. A* **57**, 4176–4187 (1998).
28. Wallraff, A. *et al.* Approaching unit visibility for control of a superconducting qubit with dispersive readout. *Phys. Rev. Lett.* **95**, 060501 (2005).

Acknowledgements This work was supported in part by the National Security Agency under the Army Research Office, by the NSF, and by Yale University. J.M.C. acknowledges support from an NSF Graduate Research Fellowship. J.K. and A.A.H. acknowledge support from Yale University via a Quantum Information and Mesoscopic Physics Fellowship. L.F. acknowledges partial support from the CNR-Istituto di Cibernetica, Pozzuoli, Italy. A.B. was supported by NSERC, CIAR and FQRNT.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.M. (johannes.majer@yale.edu) and R.J.S. (robert.schoelkopf@yale.edu).

LETTERS

Electronic and structural transitions in dense liquid sodium

Jean-Yves Raty¹, Eric Schwegler² & Stanimir A. Bonev³

At ambient conditions, the light alkali metals are free-electron-like crystals with a highly symmetric structure. However, they were found recently to exhibit unexpected complexity under pressure^{1–6}. It was predicted from theory^{1,2}—and later confirmed by experiment^{3–5}—that lithium and sodium undergo a sequence of symmetry-breaking transitions, driven by a Peierls mechanism, at high pressures. Measurements of the sodium melting curve⁶ have subsequently revealed an unprecedented (and still unexplained) pressure-induced drop in melting temperature from 1,000 K at 30 GPa down to room temperature at 120 GPa. Here we report results from *ab initio* calculations that explain the unusual melting behaviour in dense sodium. We show that molten sodium undergoes a series of pressure-induced structural and electronic transitions, analogous to those observed in solid sodium but commencing at much lower pressure in the presence of liquid disorder. As pressure is increased, liquid sodium initially evolves by assuming a more compact local structure. However, a transition to a lower-coordinated liquid takes place at a pressure of around 65 GPa, accompanied by a threefold drop in electrical conductivity. This transition is driven by the opening of a pseudogap, at the Fermi level, in the electronic density of states—an effect that has not hitherto been observed in a liquid metal. The lower-coordinated liquid emerges at high temperatures and above the stability region of a close-packed free-electron-like metal. We predict that similar exotic behaviour is possible in other materials as well.

The negative melting slope of sodium is in a pressure range in which the stable solid structures are known to be body-centred cubic (b.c.c.) and, above 65 GPa, to be the more compact face-centred cubic (f.c.c.). At around 103 GPa, solid sodium undergoes a Peierls distortion towards a cubic structure with a 16-atom cell (*cI16* space group)^{4–6}, and the melting curve recovers⁶ a positive slope. The anomalous melting behaviour of sodium suggests marked changes in the properties of the liquid that are of a different nature from what has been observed in other alkali metals previously. Indeed, the melting anomaly in sodium is one order of magnitude in temperature and two orders of magnitude in pressure larger than in caesium^{7,8}. A negative melting slope in caesium is caused by charge transfer, from *s* to more localized *d* states, which commences at a slightly lower pressure in the liquid than in the solid⁹. Because of a lack of *d* states in the valence shell, a similar mechanism is not possible in sodium. Furthermore, an analogy with the melting curve of hydrogen¹⁰, which has also been predicted to have a negative slope above that of a close-packed solid, is inappropriate as well because the unusual melting behaviour of hydrogen is explained by enhanced intermolecular charge transfer in the liquid phase.

To investigate the structural and electronic changes in compressed sodium that are responsible for the shape of its melting curve, we

carried out a series of first-principles molecular dynamic (FPMD) simulations between 5 and 120 GPa and up to 1,500 K. First, we computed the melting curve over this pressure range. Our melting calculations were carried out in the *NVT* ensemble (where *N*, *V* and *T* are the number of particles, volume and temperature, respectively), in which the solid phases are heated gradually until melting occurs. Although more accurate methods for melting computations exist (for example, two-phase simulations^{10,11} or free energy matching¹²), the ‘heat-until-it-melts’ approach used here is sufficient if there are no large overheating effects. This is the case for sodium, in which the bonding properties do not change significantly on melting and the local order in the liquid and solid phases, as shown below, is similar. Thus, melting temperatures are computed by progressively heating the three solid structures of sodium known to be stable at room temperature (300 K) up to 130 GPa, namely b.c.c., f.c.c. and *cI16* (ref. 13). The simulations are carried out with supercell sizes of 128, 108 and 128 atoms, respectively, which are sufficiently large for reproducing the sodium equation of state at room temperature⁴.

The melting curve obtained in this manner is shown in Fig. 1 and can be compared with existing experimental measurements. The

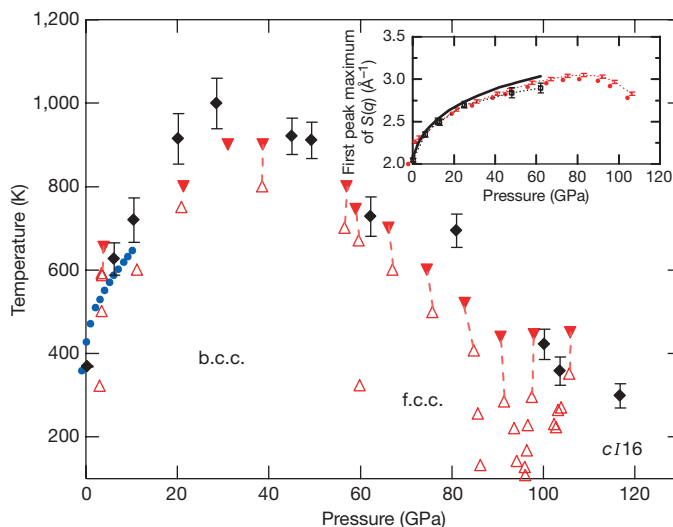


Figure 1 | The melting curve of sodium under pressure. The theoretical results are shown in red. Open uptriangles and filled downtriangles indicate simulations of solid and liquid phases, respectively. Data points of solid and liquid simulations on the same isochore that bracket the melting temperature are connected with dashed lines. Experimental measurements are shown with black diamonds⁶ and with filled blue circles¹⁴. The inset shows the evolution of the first peak maximum of $S(q)$ with pressure. Red circles are the theoretical results (error bars represent s.d.), black squares indicate measurements⁶, and the solid line is the b.c.c. (100) peak position.

¹FNRS-University of Liège, Sart-Tilman 4000, Belgium. ²Lawrence Livermore National Laboratory, Livermore, California 94550, USA. ³Department of Physics, Dalhousie University, Halifax, Nova Scotia B3H 3J5, Canada.

computed melting temperatures (T_m) are slightly lower than the experimental data in ref. 6 and are slightly higher than the data in ref. 14. The overall agreement is good and the shape of the theoretical melting curve, in particular, closely mimics the experimental data. The agreement of the computed structure factor, $S(q)$, with measurements is also excellent. From our simulations we can also estimate the finite-temperature phase boundaries between b.c.c., f.c.c. and cI16 (grey areas in Fig. 1). Simulations of any phase outside the indicated stability regions led to amorphous deformations of the structures within a few picoseconds.

The first notable feature of the melting curve is the appearance of a maximum between 20 and 40 GPa and a negative slope at higher pressures. This turnover is above the stability region of the solid b.c.c. phase, which according to the Clausius–Clapeyron equation indicates that structural changes in the liquid are taking place that make it denser than the b.c.c. solid. Such changes are not entirely surprising. In the b.c.c. phase each atom has eight nearest neighbours, whereas in the higher-pressure f.c.c. crystal each atom has 12 nearest neighbours. It is generally assumed, but rarely shown quantitatively, that molten metals exhibit local orders similar to those of their crystalline phases¹⁵. Furthermore, because finite temperature fluctuations enable structural changes in liquids to take place gradually, they may commence at lower pressures than their analogous solid–solid transitions. This opens up the possibility for a melting curve maximum if the sequence of solid phase transitions with increasing pressure is from less-compact to more-compact structures.

We therefore focused on identifying changes in the local order of the liquid structure along the melting curve that have similarities

with the b.c.c.-to-f.c.c. transition. However, comparison of such a compact yet disordered local order to an ideal crystalline structure is not straightforward. In particular, the order parameters on the basis of the characteristics of the first coordination shell, such as pair correlation functions, angular bond distributions, Steinhart's angular cluster parameters¹⁶ and individual neighbour distance distributions, are nearly invariant up to approximately 60 GPa, as was found in another recent study¹⁷.

Instead, we focused our examination on the second coordination shell. Distance distribution analysis (Fig. 2a) showed that, starting at 30 GPa, the interatomic distances in the second shell decrease more rapidly with pressure compared to those in the first shell. To interpret these changes as a progressive transition from b.c.c.-to-f.c.c.-like local order, it is helpful to compare them with the interatomic distributions of finite-temperature b.c.c. and f.c.c. solids (the atomic arrangements in b.c.c. and f.c.c. crystals are shown in Fig. 2b). Such a comparison has to be made at the same density so we used a simplified model for finite-temperature solids (Fig. 2c). The result of this comparison is that, with increasing temperature, the first coordination shells of b.c.c. and f.c.c. become almost identical. However, the second coordination shell of the heated f.c.c. is contracted compared to the second coordination shell of the heated b.c.c., in exactly the same way as it is observed in the liquid between 0 and 60 GPa. This model, together with the analysis of the FPMD simulations, demonstrates that: first, there are changes in the liquid structure, noticeable in the second coordination shell, that correlate well with the shape of

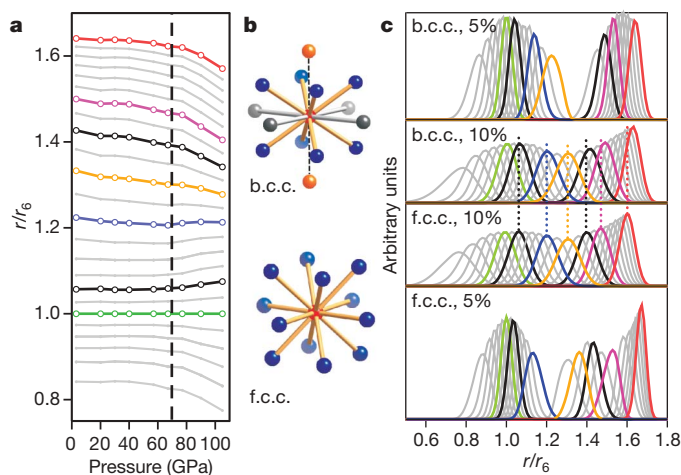


Figure 2 | Distribution of interatomic distances in sodium as a function of pressure and temperature. **a**, Pressure dependence of the average interatomic distances to the first, second, third, and so on, nearest neighbours along the melting curve in liquid sodium. Distances are scaled with respect to the sixth-nearest-neighbour distance, r_6 , to facilitate comparison at different densities, and several of the curves are highlighted with different colours. The dotted line indicates the pressure above which significant structural changes in the first coordination shell are observed. **b**, Atomic arrangements in the f.c.c. and b.c.c. structures. The f.c.c. cluster composed of the eight nearest neighbours (blue) and the four next-nearest neighbours (grey); the two remaining next-nearest neighbours in b.c.c. are shown in orange and connected with dashed lines to the central atom. **c**, Probability distributions of the interatomic distances computed for model finite temperature f.c.c. and b.c.c. solids. Thermal disorder in the model structures is introduced by displacing the atomic positions (x) according to a gaussian distribution, $\delta x = \exp(-d_1^2/2\sigma^2)$. Here d_1 is the nearest-neighbour distance in the 0 K crystal, and $\sigma = 0.1$ corresponds to the measured Debye–Waller factor in the simulated sodium system (40 GPa, 600 K). Increased disorder (expressed in per cent of the 0 K solid nearest neighbour distance) corresponds to increased temperature. The validity of this approach is illustrated in the Supplementary Information.

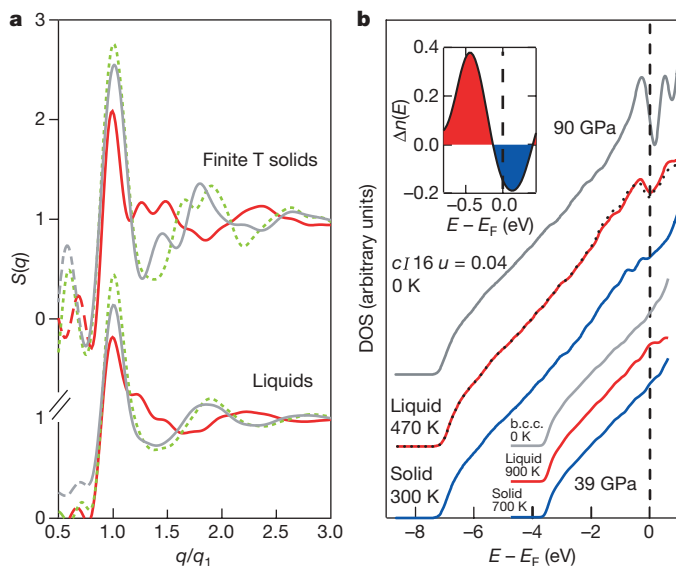


Figure 3 | Structure factors $S(q)$ and electronic DOS in solid and liquid sodium as a function of pressure and temperature. **a**, The bottom three curves are liquid $S(q)$ values computed at the following pressure and temperature conditions: 3 GPa and 650 K (grey line), 59 GPa and 745 K (dotted line) and 106 GPa and 500 K (red). The upper three curves are computed $S(q)$ values of the following solid phases: b.c.c. at 20 GPa and 700 K (grey line), f.c.c. at 69 GPa and 500 K (dotted line) and cI16 at 105 GPa and 200 K (red). To compare $S(q)$ values at different densities, the curves are scaled with respect to their first peak positions, q_1 . The $S(q)$ values are computed by Fourier transforming $(g(r) - 1)$, where the $g(r)$ values are scaled to approach unity smoothly for $r > 15$ a.u. Finite-size effects result in spurious oscillations for small q values, indicated with dashed lines in the figure. **b**, At 39 GPa (bottom three curves), the DOS of the liquid is similar to that of the heated b.c.c. solid and to the perfect b.c.c. crystal. At 90 GPa (upper three curves), the liquid DOS exhibit a pseudogap, which is not present in the f.c.c. solid at the same pressure, but is similar to the pseudogap in the 0 K cI16 crystal. The dotted line is the DOS obtained from a molecular dynamics simulation carried out with 8 k-points in the Brillouin zone; the agreement with the Γ -point only simulation is excellent. The inset shows the difference between the DOS, $n(E)$, of the liquid and of the solid near melting. Here E and E_F are energy and Fermi energy, respectively.

the melting curve; and, second, these changes are characteristic of a transition from b.c.c. to f.c.c. local order at a finite temperature.

The structural changes in the liquid to an f.c.c.-like local order, however, cannot be used to explain the anomalous melting above 60 GPa where the solid itself assumes the close-packed f.c.c. structure. Indeed, we observed significant structural changes above 65 GPa. These are already evident from the shapes of the pair correlation functions, $g(r)$ (see also Supplementary Fig. 2) and become quite clear from the neighbour distance distributions. The data in Fig. 2a, in particular, demonstrate that the first coordination shell splits and the coordination, determined by counting the neighbour-peak positions up to the first $g(r)$ minimum, drops from about 13 at 60 GPa to approximately 8 at 105 GPa. The new coordination in the liquid indicates that its local order bears similarities to a distorted b.c.c.-like structure. One such crystalline structure is *cI16*, which is the stable solid phase found at much higher pressure^{4–6}, namely above 103 GPa (for a detailed description of *cI16*, see ref. 3). In the solid, the Peierls mechanism that stabilizes *cI16* corresponds to a symmetry-breaking of b.c.c. and to the appearance of a pseudogap at the Fermi level, thus lowering the electronic energy.

To confirm the parallel between molten sodium and the *cI16* crystal, we computed the $S(q)$ of the liquid at several pressures above the melting curve and compared it with that of the finite-temperature b.c.c., f.c.c. and *cI16* solids (Fig. 3a). As pressure is increased, the changes in the liquid $S(q)$ match the variations observed over the three solid phases extremely well. Just as in the *cI16* crystal, the liquid structure factor above 65 GPa plotted against q/q_1 , where q_1 is the first peak position of $S(q)$, develops a pronounced shoulder next to the principal maximum and there is a phase change in the large q oscillations. The parallel between the liquid and solid phases is further supported by evidence in our calculations of electronic structure modifications in the liquid that drive the structural changes under pressure (Fig. 3b). We found that the electronic density of states (DOS) of liquid and solid sodium are similar at lower pressure (<40 GPa) and do not deviate noticeably from the ideal b.c.c. crystal. However, above 65 GPa, a pseudogap opens at the Fermi level in the liquid DOS that is not present in the underlying f.c.c. solid, but is similar to the pseudogap in the high-pressure *cI16* crystal. As a result, the valence states in liquid sodium are pushed towards lower energies; at 91 GPa, for instance, we measured a decrease in band energy of 40 meV per atom in the liquid compared to the f.c.c. solid (see inset in Fig. 3b). This large gain in band energy is sufficient to overcome the cost in increased repulsion associated with local distortions and

the loss in entropy compared to the non-distorted liquid. The melting temperature continues to decrease with pressure as the modifications in the liquid progress, until the f.c.c. to *cI16* solid phase transition takes place around 100 GPa. At this point, a pseudogap opens in the solid as well and the system recovers a positive slope in the melting curve.

This symmetry-breaking in molten sodium is totally counterintuitive for a high-pressure and high-temperature liquid. It is clearly related to a Peierls-like mechanism, as indicated by the appearance of a pseudogap together with the split of the first coordination shell¹⁸. A pseudogap opening by means of a Peierls-like distortion has indeed been observed in simulations of semiconductor liquids (for example, arsenic¹⁹, GeTe and GeSe²⁰), but only at ambient pressure, and in these cases it coincides with Peierls distortions in the underlying solids. Furthermore, the Peierls distortion observed in the stable crystalline phases of arsenic and antimony is known to be destroyed under pressure²¹, and in GeTe and GeSe under heating²². Any evidence for a Peierls distortion in antimony disappears on melting¹⁹.

To understand why this phenomenon occurs in sodium, we examined the effects of pressure and temperature on the angular-momentum projections of the electronic DOS. The distortion from b.c.c. to *cI16* at 0 K brings additional p states below the Fermi level (Fig. 4a, bottom). At low pressure, the very small lowering of the band structure energy is not sufficient to overcome the ionic core repulsion associated with the distortion. The effect is much more pronounced at higher pressures at which the interatomic distances are shorter and the formation of 'bonding' p states is more favourable. Increasing temperature, on the other hand, introduces disorder that results in further broadening of the p band and additional p states below the Fermi level (Fig. 4a, upper curves). This broadening is seen at finite temperature for both the solid and the liquid phases. However, the inherent disorder in the liquid enables the sampling of configurations with a local order similar to that of *cI16*.

These are general observations, which imply that similar symmetry-breaking by means of a s -to- p transfer mechanism could be expected in other liquids. It should be borne in mind, however, that the symmetry-breaking and the corresponding decrease in volume come at the expense of increased short-range ionic repulsion. Therefore, they are only feasible as long as these repulsions are 'soft' enough that they can be compensated for by the lowering of the electronic band structure energy¹⁸. This is the reason why similar effects in semiconductors are only observed in the low-pressure liquids of the lighter group-V elements, for example, in phosphorus

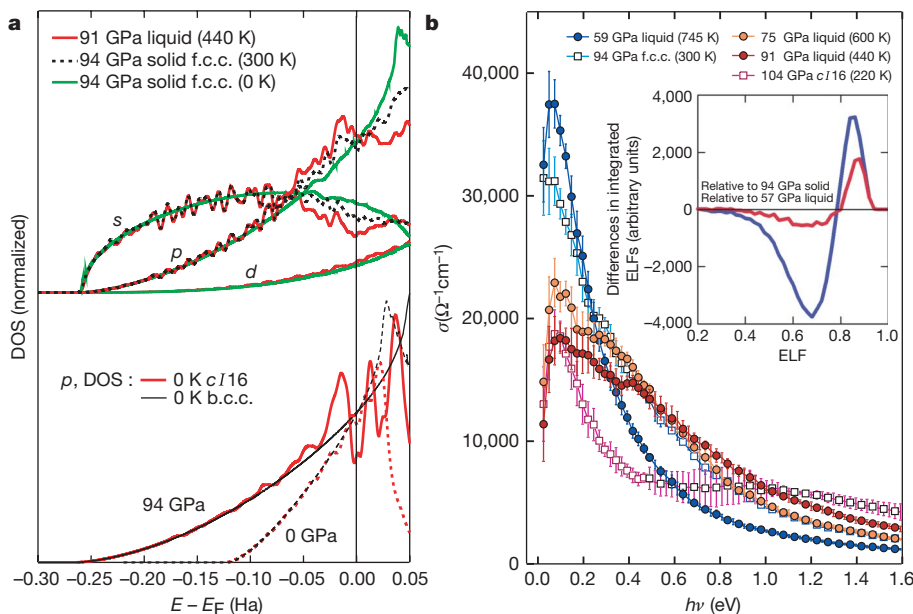


Figure 4 | Physical origin and properties of the low-coordination phase of dense liquid sodium.

a, The bottom part of the graph shows densities of p states in b.c.c. and *cI16* (with distortion parameter $u = 0.04$; where $u = 0$ results in b.c.c.) at 0 GPa (dashed lines) and 94 GPa (solid lines). The top shows s , p and d partial DOS in the 91 GPa liquid (red line) compared with the 0 K f.c.c. solid (green line) and the heated solid (black dashed line). **b**, Electrical conductivity in liquid and solid phases of sodium as a function of pressure (error bars represent s.d.). The inset shows differences (Δ) in integrated ELF²⁴. A higher ELF value indicates stronger electron localization, whereas a value of 0.5 corresponds to free-electron behaviour. There is increased localization in the 91 GPa liquid.

and arsenic, but not in antimony or bismuth^{19,23}. However, we predict that unusual liquid phases and melting behaviour is likely to be observed in other alkali and alkali-earth metals at high pressure—systems in which pressure-induced electronic instabilities produce lower-symmetry crystalline phases. There is already strong evidence that similar effects exist in lithium (I. Tamblyn, J. Y. Raty & S. A. Bonev, unpublished work).

According to these observations, the physical properties of liquid sodium are expected to change in the pressure range above 60 GPa. Indeed, the electrical conductivity computed by means of the Kubo–Greenwood formula (Fig. 4b) shows a very strong decrease, by a factor of ~ 3 between 40 and 80 GPa. The conductivity curve actually deviates from the Drude shape towards a semiconductor-like shape. Although the extrapolation to zero frequency is difficult, we estimate the direct-current conductivity to decrease from about $32,000 \Omega^{-1} \text{cm}^{-1}$ in the f.c.c. solid at 94 GPa and 300 K to less than $10,000 \Omega^{-1} \text{cm}^{-1}$ in the liquid at 470 K (91 GPa). This large drop is due to the combination of two effects, namely a decrease of the DOS and an increase in their localization near the Fermi level. To illustrate this, we have computed electron localization functions (ELFs)²⁴. The ELFs (Fig. 4b, inset) indicate stronger electron localization in the broken symmetry phase compared with both the lower-pressure liquid and the finite-temperature f.c.c. solid. We have also confirmed this by computing maximally localized Wannier functions²⁵, which exhibit decreased spread in the liquid above 65 GPa. Other unusual optical properties in this phase are likely. These findings could be confirmed by combined structure and conductivity measurements performed at modern synchrotron sources. We anticipate that our work will stimulate future experiments to confirm our predictions for sodium and to follow our suggestions for discovering similar behaviour in other dense liquid metals.

METHODS SUMMARY

The electronic structure was computed using *ab initio* density functional theory calculations²⁶. We used the Perdew, Burke and Ernzerhof generalized gradient approximation to the exchange and correlation potential and a 20 Ry plane-wave basis expansion. This level of accuracy allows for relatively fast calculations. Also, it reproduces the FLAPW (full-potential linear augmented plane wave method) b.c.c. and f.c.c. crystal equation of states⁴ up to 120 GPa within a 1 GPa error at the highest pressure, and the experimental phase boundaries between the b.c.c., f.c.c. and the *cI16* solids. The Troullier–Martins atomic pseudopotential includes the valence 3 *s* electrons and nonlinear core corrections. The molecular dynamics trajectories were integrated using our implementation of the velocity Verlet algorithm in the Quantum Espresso package (<http://www.pwscf.org>). To obtain accurate melting temperatures, the equations of motion were integrated with a time step as small as 30 atomic units (a.u.) (40 a.u. was used after melting occurred). The temperature was controlled by means of a Berendsen thermostat²⁷ with a relaxation time of 0.1 ps. The analysis of electronic properties was performed using the ABINIT package (ref. 28 and <http://www.abinit.org>).

The ‘heat-until-it-melts’ approach was effective at pressures above 25 GPa, at which melting leads to a densification of the structure. At these pressures, melting was observed within a few picoseconds. Overheating effects are possibly present below 25 GPa, because very long simulation times were necessary to melt the system (30 ps were necessary to melt the system at 20 GPa and 800 K). Therefore, the data points corresponding to the liquid phase below 30 GPa on Fig. 1 should be considered as upper boundaries for the melting temperatures.

To detect fine features of the DOS and to converge the conductivity calculations it is necessary to carry out a very careful *k*-point sampling of the Brillouin zone. Our results were obtained by performing calculations with 1,000 *k* points on 108 or 128 atomic configurations taken from the FPMD simulations. For the FPMD simulations, in contrast, a Γ -point-only sampling is sufficient; as illustrated in Fig. 3a, the resulting DOS profile is identical to that obtained from FPMD simulations carried out with an 8-*k*-point mesh. The comparison was established for two selected pressures at which fully independent molecular dynamics trajectories have been generated using either a single-*k*-point (Γ) or an 8-*k*-point sampling of the Brillouin zone. Both the structures and the electronic properties (melting temperature, local order, electronic densities of states and conductivities) were indistinguishable. All other simulations were then performed with the Γ -point-only sampling of the Brillouin zone.

Received 27 February; accepted 19 July 2007.

- Neaton, J. B. & Ashcroft, N. W. Pairing in dense lithium. *Nature* **400**, 141–144 (1999).
- Neaton, J. B. & Ashcroft, N. W. On the constitution of sodium at higher densities. *Phys. Rev. Lett.* **86**, 2830–2833 (2001).
- Hanfland, M., Syassen, K., Christensen, N. E. & Novikov, D. L. New high-pressure phases of lithium. *Nature* **408**, 174–178 (2000).
- Hanfland, M., Loa, I. & Syassen, K. Sodium under pressure: bcc to fcc structural transition and pressure–volume relation to 100 GPa. *Phys. Rev. B* **65**, 184109 (2002).
- Syassen, K. in *High-Pressure Phenomena* (eds Hemley, R. J., Chiarotti, G., Bernasconi, M. & Ulivi, L.) 251–271 (IOS, Amsterdam, 2002).
- Gregoryanz, E., Degtyareva, O., Somayazulu, M., Hemley, R. J. & Mao, H. Melting of dense sodium. *Phys. Rev. Lett.* **94**, 185502–1–4 (2005).
- Jayaraman, A., Newton, R. C. & McDonough, J. M. Phase relations, resistivity, and electronic structure of cesium at high pressures. *Phys. Rev.* **159**, 527–533 (1967).
- Boehler, R. & Zha, C.-S. Systematics in the melting behavior of the alkali metals from DAC measurements. *Physica B* **139–140**, 233–236 (1986).
- Glözel, D. & McMahon, A. K. Relativistic effects, phonons, and the isostructural transition in cesium. *Phys. Rev. B* **20**, 3210–3216 (1979).
- Bonev, S. A., Schwegler, E., Ogitsu, T. & Galli, G. A quantum fluid of metallic hydrogen suggested by first-principles calculations. *Nature* **431**, 669–672 (2004).
- Correa, A., Bonev, S. A. & Galli, G. Carbon under extreme conditions: phase boundaries and electronic properties from first-principles theory. *Proc. Natl Acad. Sci. USA* **103**, 1204–1208 (2006).
- Sugino, O. & Car, R. *Ab initio* molecular dynamics study of first-order phase transitions: melting of silicon. *Phys. Rev. Lett.* **74**, 1823–1826 (1995).
- Ackland, G. J. & McLeod, I. R. Origin of the complex crystal structures of elements at intermediate pressure. *New J. Phys.* **6**, 138–140 (2004).
- Zha, C.-S. & Boehler, R. Melting of sodium and potassium in a diamond anvil cell. *Phys. Rev. B* **31**, 3199–3201 (1985).
- Waseda, Y. *The Structure of Non-Crystalline Materials: Liquids and Amorphous Solids* ch. 3 6 (McGraw Hill, Texas, 1980).
- Steinhardt, P. J., Nelson, D. R. & Ronchetti, M. Icosahedral bond orientational order in supercooled liquids. *Phys. Rev. Lett.* **47**, 1297–1300 (1981).
- Hernandez, E. R. & Iniguez, J. First principles simulations on the nature of the melting line of sodium. *Phys. Rev. Lett.* **98**, 055501–1–4 (2007).
- Gaspard, J.-P., Pellegatti, A., Marinelli, F. & Bichara, C. Peierls instabilities in covalent structures I. Electronic structure, cohesion and the $Z = 8 - N$ rule. *Phil. Mag.* **77**, 727–744 (1998).
- Bichara, C., Pellegatti, A. & Gaspard, J. P. Properties of liquid group-V elements: a numerical tight-binding simulation. *Phys. Rev. B* **47**, 5002–5007 (1993).
- Raty, J. Y. *et al.* Distance correlations and dynamics of liquid GeSe: an *ab initio* molecular dynamics study. *Phys. Rev. B* **64**, 235209–1–7 (2001).
- Beister, H. J., Strössner, K. & Syassen, K. Rhombohedral to simple-cubic phase transition in arsenic under pressure. *Phys. Rev. B* **41**, 5535–5543 (1990).
- Chattopadhyay, T., Boucherle, J. X. & Von Schnering, H. G. Neutron diffraction study on the structural phase transition in GeTe. *J. Phys. C* **20**, 1431–1440 (1987).
- Katayama, Y. & Tsuji, K. X-ray structural studies on elemental liquids under high pressures. *J. Phys. Condens. Matter* **15**, 6085–6103 (2003).
- Becke, A. D. & Edgecombe, K. E. A simple measure of electron localization in atomic and molecular systems. *J. Chem. Phys.* **92**, 5397–5403 (1990).
- Marzari, N. & Vanderbilt, D. Maximally localized generalized Wannier functions for composite energy bands. *Phys. Rev. B* **56**, 12847–12865 (1997).
- Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
- Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
- Gonze, X. *et al.* First principles computation of material properties: the ABINIT software project. *Comp. Mat. Sci.* **25**, 478–492 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank I. Souza and N. W. Ashcroft for discussions. This work was supported by the NSERC of Canada. J.-Y.R. acknowledges support by the FNRS and the FAME NoE. E.S. worked under the auspices of the US Department of Energy at the University of California/Lawrence Livermore National Laboratory.

Author Contributions J.-Y.R., E.S. and S.A.B. contributed equally to this work. J.-Y.R. and S.A.B. designed the research. J.-Y.R. and E.S. conducted the molecular dynamics simulations. J.-Y.R., E.S. and S.A.B. performed the data analysis (S.A.B. and J.-Y.R. computed the densities of states and conductivities; E.S. performed the Wannier analysis; J.-Y.R. performed the model calculations and the ELF analysis; and S.A.B. performed the solid state calculations). J.-Y.R. and S.A.B. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.-Y.R. (jyraty@ulg.ac.be) or S.A.B. (stanimir.bonev@dal.ca).

LETTERS

Millennial-scale trends in west Pacific warm pool hydrology since the Last Glacial Maximum

Judson W. Partin¹, Kim M. Cobb¹, Jess F. Adkins², Brian Clark³ & Diego P. Fernandez²

Models and palaeoclimate data suggest that the tropical Pacific climate system plays a key part in the mechanisms underlying orbital-scale and abrupt climate change^{1–7}. Atmospheric convection over the western tropical Pacific is a major source of heat and moisture to extratropical regions, and may therefore influence the global climate response to a variety of forcing factors. The response of tropical Pacific convection to changes in global climate boundary conditions, abrupt climate changes and radiative forcing remains uncertain, however. Here we present three absolutely dated oxygen isotope records from stalagmites in northern Borneo that reflect changes in west Pacific warm pool hydrology over the past 27,000 years. Our results suggest that convection over the western tropical Pacific weakened 18,000–20,000 years ago, as tropical Pacific^{2,5,6,8} and Antarctic⁹ temperatures began to rise during the early stages of deglaciation. Convective activity, as inferred from oxygen isotopes, reached a minimum during Heinrich event 1 (ref. 10), when the Atlantic meridional overturning circulation was weak¹¹, pointing to feedbacks between the strength of the overturning circulation and tropical Pacific hydrology. There is no evidence of the Younger Dryas event¹² in the stalagmite records, however, suggesting that different mechanisms operated during these two abrupt deglacial climate events. During the Holocene epoch, convective activity appears to track changes in spring and autumn insolation, highlighting the sensitivity of tropical Pacific convection to external radiative forcing. Together, these findings demonstrate that the tropical Pacific hydrological cycle is sensitive to high-latitude climate processes in both hemispheres, as well as to external radiative forcing, and that it may have a central role in abrupt climate change events.

Numerous palaeoclimatic studies have focused on changes in the tropical Pacific zonal sea surface temperature (SST) gradient and associated shifts in convection^{2,4,5}, extending an El Niño/Southern Oscillation (ENSO) framework to interpretations of millennial-scale climate variability. Recently, the potential for meridional changes in tropical Pacific temperatures has emerged as an important mechanism in shaping hydrological responses to abrupt climate change^{7,13–15}. Indeed, new modelling and palaeoclimate results suggest that during Heinrich event 1 (H1), when a near-collapse of the Atlantic meridional overturning circulation¹¹ cooled most of the Northern Hemisphere, the Intertropical Convergence Zone (ITCZ) shifted southwards^{7,13–15}. However, the extent to which tropical Pacific climate feedbacks were involved in deglacial abrupt climate events such as H1, the Antarctic Cold Reversal⁹ and subsequent Younger Dryas¹² remains unclear. A combination of zonal and/or meridional changes in the distribution of tropical convection probably played a key part in shaping the global climate response to abrupt climate changes during the deglaciation. Uncovering the

mechanisms and feedbacks that govern the response of the tropical Pacific climate system to both internal and external forcings requires well-dated, high-resolution records from climatic centres of action.

Here we present three absolutely dated stalagmite oxygen isotopic ($\delta^{18}\text{O}$) records from northern Borneo that document changes in western tropical Pacific atmospheric circulation and hydrology over the past 27,000 yr. Tropical stalagmite $\delta^{18}\text{O}$ records are particularly well-suited to the investigation of centennial-to-millennial hydrological change because tropical rainfall $\delta^{18}\text{O}$ is inversely correlated to precipitation amount and because U–Th dating provides excellent chronological control.

The research site is located in Gunung Buda National Park (4° N, 114° E) in the northwestern corner of Malaysian Borneo (see Methods for detailed site description). The ITCZ lies above northern Borneo year-round, delivering 5 m of rainfall with little seasonality (Supplementary Fig. 1). ENSO exerts a dominant control on northern Borneo precipitation, with anomalies as large as $\pm 50\%$ during ENSO extremes (Fig. 1). Interannual (2–7 yr) changes in rainfall account for $\sim 20\%$ of total precipitation variance in northern Borneo and are highly correlated to the Southern Oscillation Index ($R = -0.80$). Seasonal cycles in rainwater $\delta^{18}\text{O}$ (-10% during boreal autumn and -4% during boreal spring) probably reflect northern versus southern moisture trajectories¹⁶. ENSO-related interannual rainfall $\delta^{18}\text{O}$ variability is consistent with the ‘amount effect’^{17,18}, whereby periods of increased precipitation are characterized by lighter rainfall $\delta^{18}\text{O}$ (ref. 16).

Several tests confirm that the Gunung Buda stalagmites formed under oxygen isotopic equilibrium, allowing carbonate $\delta^{18}\text{O}$ changes

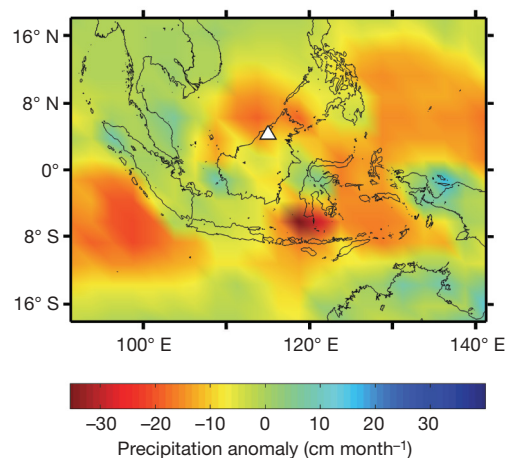


Figure 1 | Map of the west Pacific December–January–February precipitation anomaly during the 1997–98 El Niño event. Data are from ref. 31. A white triangle marks the approximate location of the research site.

¹School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, USA. ²Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA. ³Gunung Mulu National Park, Sarawak, Malaysia.

to be interpreted as rainwater $\delta^{18}\text{O}$ changes. First, the correlation between oxygen and carbon isotopic variability is low ($R < 0.05$) after linearly detrending for glacial–interglacial isotopic changes. Second, all three samples pass the ‘Hendy test’¹⁹, exhibiting $\delta^{18}\text{O}$ variations of less than $\sim 0.5\text{‰}$ across the axial portion of a single growth layer (see Methods and Supplementary Figs 9 and 10). Third, the equilibrium calcite $\delta^{18}\text{O}$ value equals that measured for modern stalagmite calcite, given measured present-day rainwater $\delta^{18}\text{O}$ of -6.8‰ and 26°C cave temperatures¹⁶. Last, the high degree of millennial-scale reproducibility of $\delta^{18}\text{O}$ in the three stalagmites from caves located $\sim 5\text{ km}$ apart strongly suggests that the stalagmite $\delta^{18}\text{O}$ records regional climate changes associated with rainfall $\delta^{18}\text{O}$ variability (Fig. 2). Poor sub-millennial reproducibility between the three records can be attributed to dating uncertainties and/or site-specific stalagmite $\delta^{18}\text{O}$ variability. It is important to note that temperature changes could only account for $\sim 0.7\text{‰}$ of stalagmite $\delta^{18}\text{O}$ variability over the past 27,000 yr, given that warm pool temperatures were no more than 3.5°C colder during the Last Glacial Maximum (LGM)^{2,5,6,8}.

A total of 75 U–Th dates and 11 isochrons provide excellent chronological control for the three stalagmite $\delta^{18}\text{O}$ records (see Methods and Supplementary Information). The age model for each record consists of 24–26 U–Th dates (Fig. 2) that were corrected for detrital thorium using stalagmite-specific detrital $^{230}\text{Th}/^{232}\text{Th}$ values calculated using isochrons. Age errors of up to 2% (2σ) represent a combination of analytical uncertainty and uncertainty in the detrital ^{232}Th correction. Slow growth rates and/or unresolved hiatuses represent the largest sources of chronological uncertainty, so we limit our climatic interpretations to portions of the records with growth rates higher than $10\text{ }\mu\text{m yr}^{-1}$.

LGM stalagmite $\delta^{18}\text{O}$ values (averaged over the period 19–23 kyr ago) are $1.3 \pm 0.3\text{‰}$ heavier than modern values, and reflect a combination of global ice volume ($+1\text{‰}$)²⁰, LGM cooling of $\sim 2\text{--}3.5^\circ\text{C}$ in the western tropical Pacific^{2,5,6,8} ($+0.4\text{‰}$ to $+0.7\text{‰}$), and poorly constrained changes in LGM regional seawater $\delta^{18}\text{O}$ (-0.5‰ to $+0.5\text{‰}$)^{2,5,6,21}. Thus, potentially small changes in northern Borneo

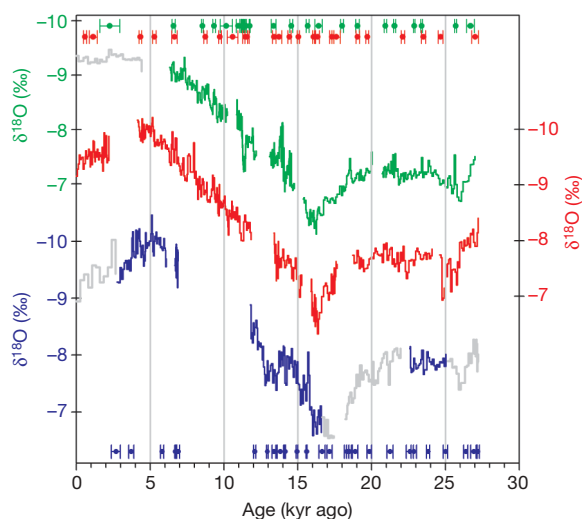


Figure 2 | Three absolutely dated stalagmite $\delta^{18}\text{O}$ records from northern Borneo. U–Th dates for each record are plotted in corresponding colours along the top and bottom; error bars represent 2σ analytical uncertainty plus uncertainty in the detrital thorium correction. The average $\delta^{18}\text{O}$ temporal resolution is 72, 56 and 60 yr per sample for SCH02, SSC01 and BA04, respectively, but varies from 1 to 100 yr per sample depending on growth rate. Data depicted in grey represent slow-growing ($<10\text{ }\mu\text{m yr}^{-1}$) portions of the records, and are considered untrustworthy for climatic interpretation owing to poor chronological control (up to $\pm 1\text{ kyr}$). Sample BA04 grew in Bukit Assam cave, while SSC01 and SCH02 grew $\sim 20\text{ m}$ apart in Snail Shell cave, roughly 6 km from Bukit Assam (see Methods for more detailed cave and sample descriptions).

rainfall $\delta^{18}\text{O}$ during the LGM are difficult to resolve. Possible controls on northern Borneo rainfall $\delta^{18}\text{O}$ over our whole 27-kyr record include: (1) changes in the tropical Pacific zonal SST gradient; (2) changes in the location and/or intensity of the ITCZ; and (3) changes in eustatic sea level, which determine the size of the emergent Sunda Shelf. Exposure of the Sunda Shelf undoubtedly altered atmospheric circulation in the west Pacific by increasing continentality and lengthening moisture trajectories, both of which deplete rainfall $\delta^{18}\text{O}$ (refs 17, 18). A complete understanding of warm pool hydrological changes during the LGM and deglaciation must account for the effect of the Sunda Shelf on the tropical Pacific coupled system.

Stalagmite $\delta^{18}\text{O}$ values begin a protracted trend towards more positive values 18–20 kyr ago, as tropical Pacific SST^{2,5,6,8} and Antarctic⁹ temperatures began to rise during the early deglaciation. This stalagmite $\delta^{18}\text{O}$ excursion cannot be attributed to temperature or local seawater $\delta^{18}\text{O}$ changes, as recorded in nearby marine sediments^{5,6,8,21}, and is therefore interpreted as a positive rainfall $\delta^{18}\text{O}$ anomaly (dry conditions based on the amount effect). The inferred trend towards drier conditions in northern Borneo culminates in maximum stalagmite $\delta^{18}\text{O}$ values $16.3 \pm 0.3\text{ kyr}$ ago, coincident with the timing of a $\delta^{18}\text{O}$ maximum in a Chinese stalagmite²² attributed to H1 (Fig. 3). Conservative age error bars for specific features of the Borneo stalagmite $\delta^{18}\text{O}$ records take into account uncertainties associated with stalagmite $\delta^{18}\text{O}$ reproducibility, U-series dating errors, and the potential for nonlinear growth rates. Dry conditions in northern Borneo during H1 are consistent with model results¹⁵ indicating a southward shift of the ITCZ in conjunction with a weakened Atlantic meridional overturning circulation inferred from proxy data¹¹. Together with evidence for dry conditions in southeast

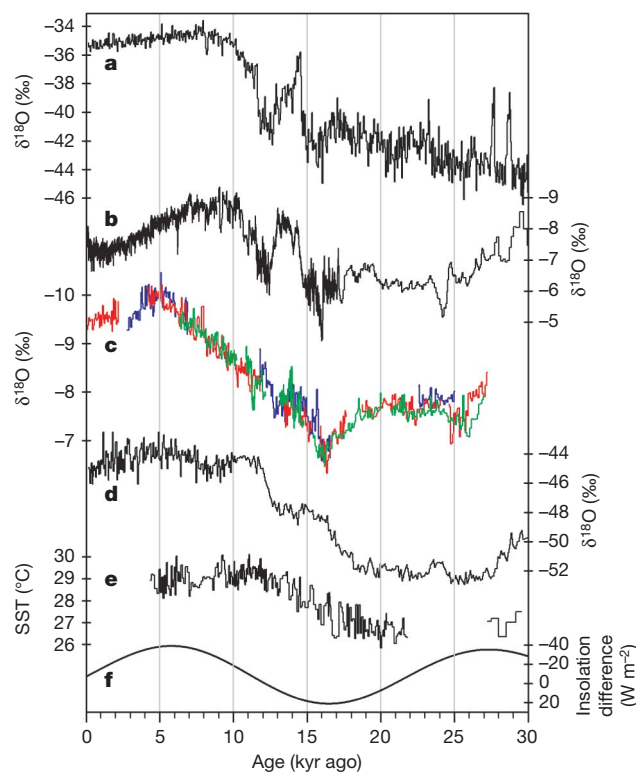


Figure 3 | Comparison of the Borneo stalagmite $\delta^{18}\text{O}$ records with other palaeoclimate records. Records plotted are as follows. **a**, Greenland (NGRIP) ice core $\delta^{18}\text{O}$ (ref. 12). **b**, Hulu/Dongge caves stalagmite $\delta^{18}\text{O}$ records (ref. 22 and R. L. Edwards, personal communication). **c**, Borneo stalagmite $\delta^{18}\text{O}$ records (SCH02, blue; SSC01, red; BA04, green). Slow-growing ($<10\text{ }\mu\text{m yr}^{-1}$) portions of the records are excluded. Note that BA04 $\delta^{18}\text{O}$ values have been shifted by $+0.4\text{‰}$. **d**, EPICA Dronning Maud Land ice core $\delta^{18}\text{O}$ (ref. 9). **e**, Sediment core reconstruction of SST from the Sulu Sea⁶. **f**, March minus September insolation at the Equator.

Asia²² and enhanced wind-driven upwelling in the eastern equatorial Pacific⁷ during H1, the new stalagmite data strongly suggest that a meridional ITCZ shift may have characterized the tropical Pacific hydrological response during H1.

However, the new stalagmite data suggest that the relationship between North Atlantic and Pacific climate is more complex than unidirectional forcing and associated response. For example, the onset of dry conditions in northern Borneo lies between 18 and 20 kyr ago, before the beginning of H1 in North Atlantic sediments¹⁰. Interestingly, this early onset of dry conditions in northern Borneo occurs during a transition to relatively wet conditions recorded in the Chinese stalagmites²², before the strong drying observed ~16.5 kyr ago in both stalagmite reconstructions. Furthermore, while sediment reconstructions of the strength of the Atlantic meridional overturning circulation suggest that a rapid resumption of overturning occurred 14.7 kyr ago¹¹, coincident with the onset of the Bølling–Allerød, the Borneo stalagmite $\delta^{18}\text{O}$ records exhibit a more gradual recovery that began as early as 15.5 kyr ago. Therefore, climatic feedbacks in the tropical Pacific may have played a part in driving the variability of the North Atlantic meridional overturning circulation across H1, which in turn affected tropical Pacific hydrology.

Following H1, the deglaciation follows a relatively smooth trend towards more negative Holocene values, interrupted by a millennium-long $\delta^{18}\text{O}$ plateau centred at 13.2 ± 0.2 kyr ago that coincides with the Antarctic Cold Reversal⁹ (Fig. 3). A western tropical Pacific expression of the Antarctic Cold Reversal is somewhat surprising, given that a prominent Younger Dryas event is present in a Chinese stalagmite record²². Oxygen isotopic records from several western tropical Pacific sediment cores north of the Equator exhibit a muted Greenland-like sequence of late deglacial events^{6,23}, including the Younger Dryas, consistent with Northern Hemisphere influence on the southeast Asian monsoon during the deglaciation. However, with one exception²³, temperature proxy records from northern west Pacific sediment cores do not contain the Younger Dryas^{6,24}, raising the prospect that the hydrological signature of the Younger Dryas in these cores is associated with runoff from the southeast Asian landmass rather than equatorial rainfall. Indeed, $\delta^{18}\text{O}$ and temperature proxy records from cores south of the Equator depict a smooth, uninterrupted deglaciation^{8,24}, reminiscent of Southern Hemisphere ice cores and the Borneo stalagmite $\delta^{18}\text{O}$ data. The similarity between the Antarctic ice core record and the western tropical Pacific stalagmite record during the late deglacial strengthens the view that Southern Hemisphere forcing dominated tropical Pacific climate during this period^{2,4,8}.

The Holocene portions of the records are characterized by broad $\delta^{18}\text{O}$ minima 5 kyr ago, indicating the sensitivity of the western tropical Pacific to spring/autumn precessional insolation forcing (Fig. 3). The stalagmite $\delta^{18}\text{O}$ minimum is interpreted as a negative rainfall $\delta^{18}\text{O}$ anomaly (wet conditions), as it cannot be explained by regional temperature and/or seawater $\delta^{18}\text{O}$ changes^{5,8,21,24}. Most tropical palaeo-precipitation records from north of the Equator contain maxima in the range 10 to 8 kyr ago^{24–26}, while those from south of the Equator exhibit late Holocene maxima²⁷, presumably linked to boreal and austral summer insolation, respectively. When combined with the new data from the Borneo stalagmites, these records strongly suggest that the mean position of the ITCZ migrated southwards over the course of the Holocene in response to precessional forcing, crossing the equatorial west Pacific ~5 kyr ago. However, relatively wet conditions in northern Borneo during the mid-Holocene could also reflect an increase in the tropical Pacific's zonal SST gradient driven by spring/autumn insolation, a prospect that finds qualitative support in coupled model simulations²⁸. Indeed, an observed minimum ~5 kyr ago in eastern equatorial Pacific temperatures⁴ lends further support to a zonally mediated insolation response, whereby an increased tropical Pacific zonal SST gradient led to enhanced warm pool convection. A minimum in atmospheric CH_4 5 kyr ago²⁹ may

provide further clues—precipitation anomalies associated with ITCZ variability versus Walker circulation changes probably have different consequences for global methane production. Increased methane production during modern-day El Niño events³⁰ suggests that a mid-Holocene increase in the tropical Pacific zonal SST gradient may have contributed to the mid-Holocene minimum in atmospheric methane. However, an ITCZ-related mechanism for atmospheric methane control during the Holocene cannot be ruled out, and warrants further investigation.

This study demonstrates that the tropical Pacific hydrological cycle is sensitive to high-latitude climate processes in both hemispheres as well as to external radiative forcing. However, the relatively smooth character of the warm pool's hydrological variability over the past 27,000 yr suggests a limited potential for large, abrupt changes in the character of tropical Pacific variability. Nonetheless, by gradually altering the heat and salt budgets of the global oceans, the tropical Pacific may have a pivotal role in driving thermohaline circulation changes associated with abrupt climate change events. Whether the tropical Pacific coupled system acts as an amplifier or a trigger of internal global climate variability, its feedbacks on the global climate system must be an integral part of any climate change mechanism, natural or anthropogenic.

METHODS SUMMARY

U-series samples weighing 100–500 mg were drilled with a 1.6 mm drill bit parallel to growth banding. Each sample was dissolved and spiked with a mixed ^{236}U – ^{229}Th solution before the separation of U and Th using standard techniques. The isotopic compositions of the U and Th fractions were measured using a Finnigan 'Neptune' MC-ICPMS at Caltech (see Methods and Supplementary Information). Relatively low ^{238}U concentrations (0.1–0.5 p.p.m.) combined with low $\delta^{234}\text{U}$ values (–100‰ to –600‰) and high detrital ^{232}Th concentrations result in U–Th age uncertainties of 0.3–2% (2σ) (see Supplementary Information). We account for initial ^{230}Th using the measured ^{232}Th content and estimates of detrital $^{230}\text{Th}/^{232}\text{Th}$ ratios ($\sim(60\text{--}120) \times 10^{-6}$ atomic ratio) obtained from 11 isochrons. All three stalagmites contain multiple hiatuses that are associated with visible bands and high ^{232}Th concentrations. Larger hiatuses are closely bracketed by U-series dates, but shorter and/or adjacent hiatuses can not be accurately resolved with U-series sampling. Slow-growing ($<10 \mu\text{m yr}^{-1}$) portions of the stalagmites probably contain unresolved hiatuses (supported by high ^{232}Th concentrations measured in these regions), and reflect large chronological uncertainties (up to 1,000 yr in the case that a hiatus is poorly constrained by high- ^{232}Th dates). Stalagmite $\delta^{18}\text{O}$ data were assigned calendar ages based on a linear interpolation between each pair of U-series dates.

Oxygen isotopic analyses were conducted on powders drilled every 1 mm along the central growth axis of the stalagmites with a 1.6 mm drill bit. The $\delta^{18}\text{O}$ of BA04 and SSC01 powders were measured using a GV Isoprime-Multiprep located at Georgia Tech (long-term reproducibility of $\pm 0.05\text{‰}$). SCH02 $\delta^{18}\text{O}$ profiles were analysed on a Finnigan 253 equipped with a Kiel device located at Wood's Hole Oceanographic Institute (long-term reproducibility of $\pm 0.08\text{‰}$). All $\delta^{18}\text{O}$ data are reported with respect to VPDB.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 February; accepted 8 August 2007.

1. Bush, A. B. G. & Philander, S. G. H. The role of ocean-atmosphere interactions in tropical cooling during the last glacial maximum. *Science* **279**, 1341–1344 (1998).
2. Lea, D. W., Pak, D. K. & Spero, H. J. Climate impact of late quaternary equatorial Pacific sea surface temperature variations. *Science* **289**, 1719–1724 (2000).
3. Clement, A. C., Cane, M. A. & Seager, R. An orbitally driven tropical source for abrupt climate change. *J. Clim.* **14**, 2369–2375 (2001).
4. Koutavas, A., Lynch-Stieglitz, J., Marchitto, T. M. & Sachs, J. P. El Niño-like pattern in ice age tropical Pacific sea surface temperature. *Science* **297**, 226–230 (2002).
5. Stott, L., Poulson, C., Lund, S. & Thunell, R. Super ENSO and global climate oscillations at millennial time scales. *Science* **297**, 222–226 (2002).
6. Rosenthal, Y., Oppo, D. W. & Linsley, B. K. The amplitude and phasing of climate change during the last deglaciation in the Sulu Sea, western equatorial Pacific. *Geophys. Res. Lett.* **30**, doi:10.1029/2002GL016612 (2003).
7. Kienast, M. et al. Eastern Pacific cooling and Atlantic overturning circulation during the last deglaciation. *Nature* **443**, 846–849 (2006).
8. Visser, K., Thunell, R. & Stott, L. Magnitude and timing of temperature change in the Indo-Pacific warm pool during deglaciation. *Nature* **421**, 152–155 (2003).

9. EPICA Community Members. One-to-one coupling of glacial climate variability in Greenland and Antarctica. *Nature* **444**, 195–198 (2006).
10. Hemming, S. R. Heinrich events: Massive late Pleistocene detritus layers of the North Atlantic and their global climate imprint. *Rev. Geophys.* **42**, doi:10.1029/2003RG000128 (2004).
11. McManus, J. F., Francois, R., Gherardi, J. M., Keigwin, L. D. & Brown-Leger, S. Collapse and rapid resumption of Atlantic meridional circulation linked to deglacial climate changes. *Nature* **428**, 834–837 (2004).
12. Rasmussen, S. O. *et al.* A new Greenland ice core chronology for the last glacial termination. *J. Geophys. Res.* **111**, doi:10.1029/2005JD006079 (2006).
13. Peterson, L. C., Haug, G. H., Hughen, K. A. & Rohl, U. Rapid changes in the hydrologic cycle of the tropical Atlantic during the last glacial. *Science* **290**, 1947–1951 (2000).
14. Chiang, J. C. H. & Bitz, C. M. Influence of high latitude ice cover on the marine Intertropical Convergence Zone. *Clim. Dyn.* **25**, 477–496 (2005).
15. Zhang, R. & Delworth, T. L. Simulated tropical response to a substantial weakening of the Atlantic thermohaline circulation. *J. Clim.* **18**, 1853–1860 (2005).
16. Cobb, K. M., Adkins, J. F., Partin, J. W. & Clark, B. Regional-scale climate influences on temporal variations of rainwater and cave dripwater oxygen isotopes in northern Borneo. *Earth Planet. Sci. Lett.* (in the press).
17. Dansgaard, W. Stable isotopes in precipitation. *Tellus* **16**, 436–468 (1964).
18. Rozanski, K., Araguas-Araguas, L. & Gonfiantini, R. in *Climate Change in Continental Isotopic Records* (eds Swart, P. K., Lohmann, K. C., McKenzie, J. & Savin, S.) 1–36 (Geophysical Monograph 78, American Geophysical Union, Washington DC, 1993).
19. Hendy, C. H. Isotopic geochemistry of speleothems. 1. Calculation of effects of different modes of formation on isotopic composition of speleothems and their applicability as palaeoclimatic indicators. *Geochim. Cosmochim. Acta* **35**, 801–824 (1971).
20. Schrag, D. P. *et al.* The oxygen isotopic composition of seawater during the Last Glacial Maximum. *Quat. Sci. Rev.* **21**, 331–342 (2002).
21. Steinke, S. *et al.* On the influence of sea level and monsoon climate on the southern South China Sea freshwater budget over the last 22,000 years. *Quat. Sci. Rev.* **25**, 1475–1488 (2006).
22. Wang, Y. J. *et al.* A high-resolution absolute-dated Late Pleistocene monsoon record from Hulu Cave, China. *Science* **294**, 2345–2348 (2001).
23. Kienast, M., Steinke, S., Stattegger, K. & Calvert, S. E. Synchronous tropical South China Sea SST change and Greenland warming during deglaciation. *Science* **291**, 2132–2134 (2001).
24. Stott, L. *et al.* Decline of surface temperature and salinity in the western tropical Pacific Ocean in the Holocene epoch. *Nature* **431**, 56–59 (2004).
25. Haug, G. H., Hughen, K. A., Sigman, D. M., Peterson, L. C. & Rohl, U. Southward migration of the intertropical convergence zone through the Holocene. *Science* **293**, 1304–1308 (2001).
26. Fleitmann, D. *et al.* Holocene forcing of the Indian monsoon recorded in a stalagmite from Southern Oman. *Science* **300**, 1737–1739 (2003).
27. Wang, X. *et al.* Interhemispheric anti-phasing of rainfall during the last glacial period. *Quat. Sci. Rev.* **25**, 3391–3403 (2006).
28. Clement, A. C., Seager, R. & Cane, M. A. Orbital controls on the El Niño/Southern Oscillation and the tropical climate. *Paleoceanography* **14**, 441–456 (1999).
29. Chappellaz, J. *et al.* Changes in the atmospheric CH₄ gradient between Greenland and Antarctica during the Holocene. *J. Geophys. Res. Atmos.* **102**, 15987–15997 (1997).
30. Dlugokencky, E. J., Walter, B. P., Masarie, K. A., Lang, P. M. & Kasischke, E. S. Measurements of an anomalous global methane increase during 1998. *Geophys. Res. Lett.* **28**, 499–502 (2001).
31. Xie, P. P. & Arkin, P. A. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Am. Meteorol. Soc.* **78**, 2539–2558 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Malang, J. Gau and S. Clark of Gunung Mulu National Park and J. Baei Hassan of Logan Bunut National Park for field assistance. J. Despain, G. Prest, S. Fryer, J. Mosenfelder and B. Hacker provided field assistance during the 2003 field trip. A. A. Tuen (UNIMAS) greatly facilitated our 2006 fieldwork in Sarawak. We also thank D. Lund for assistance in U–Th dating, and J. Lynch-Stieglitz and M. Schmidt for providing comments on early versions of the manuscript. The research was funded by NSF-ESH and by a Comer Abrupt Climate Change Fellowship.

Author Information The stalagmite $\delta^{18}\text{O}$ data can be downloaded at <http://www.ncdc.noaa.gov/paleo/pubs/partin2007/partin2007.html>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.W.P. (jpartin@eas.gatech.edu).

METHODS

U-series dating. U-series dating of carbonates has been an important part of constraining past climate for decades. With the advance of U and Th isotope detection by mass spectrometry instead of alpha particle counting, sample sizes were shrunk while also improving precision³². Stalagmites tend to have lower ^{238}U concentrations than corals but higher $^{234}\text{U}/^{238}\text{U}$ ratios³³. Our tropical stalagmites have low ^{238}U , low $^{234}\text{U}/^{238}\text{U}$ ratios and relatively high detrital ^{232}Th . These features make radiogenic age constraints especially challenging in our samples. Sample BA04 has 1–2 p.p.m. ^{238}U , about -600‰ $\delta^{234}\text{U}$, and 1–100 pmol g⁻¹ ^{232}Th . Sample SCH02 has 0.3–0.9 p.p.m. ^{238}U , about -350‰ $\delta^{234}\text{U}$, and 1–15 pmol g⁻¹ ^{232}Th . Sample SSC01 has 0.1–0.3 p.p.m. ^{238}U , about -100‰ $\delta^{234}\text{U}$, and 0.3–3 pmol g⁻¹ ^{232}Th .

We drilled 100–500 mg samples from time synchronous bands along the growth axis of all three stalagmites. Samples were dissolved and spiked with a mixed ^{236}U and ^{229}Th solution³⁴. U and Th isotopes were separated and purified by traditional methods³⁵ and measured on a Finnigan ‘Neptune’ MC-ICP-MS at Caltech. The U fraction was tested for the U concentration and diluted to match intensities of all samples. ^{234}U was measured on the centre SEM with ^{235}U and ^{238}U on separate Faraday cups. ^{229}Th and ^{230}Th were measured on the MICs attached to Faraday cup L4 and ^{232}Th was measured in this cup. For both U and Th, samples were bracketed with known ratio standards, CRM-145 for U and an in-house spiked gravimetric standard for Th. With samples of $\sim 100\text{ ng }^{238}\text{U}$ we can measure the $^{229}\text{Th}/^{230}\text{Th}$ ratio to 1–2‰ (2 σ) and the $^{234}\text{U}/^{238}\text{U}$ ratio to better than 0.5‰.

Hendy tests. Powders were drilled along a single growth layer to measure the $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ variability as a function of distance from the stalagmites’ central growth axes, in order to quantify potential effects of kinetic fractionation¹⁹. Isotopic values are plotted as departures from central axes values in Supplementary Fig. 9. Negative distances correspond to samples drilled to the left of the central axis, and positive distances represent samples drilled to the right of the central axis. Non-equilibrium CO_2 degassing on the surface of the stalagmite would result in kinetic fractionation of oxygen and carbon isotopes and isotopic enrichments off-axis. Such kinetic fractionation might explain the weak correlation between $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ variability across single growth layers of the Borneo stalagmites (Supplementary Fig. 10). However, our Hendy analyses suggest that kinetic fractionation accounts for no more than $\sim 0.5\text{‰}$ of the $\delta^{18}\text{O}$ variability observed in the down-core stalagmite $\delta^{18}\text{O}$ records. Furthermore, the fact that the three Borneo stalagmite $\delta^{18}\text{O}$ records exhibit strong millennial-scale reproducibility confirms that rainfall $\delta^{18}\text{O}$ variations are the likely source of their shared millennial-scale $\delta^{18}\text{O}$ variability.

Cave locations. Samples SSC01 and SCH02 originate from Snail Shell cave ($4^\circ 12' 20.8''\text{N}$, $114^\circ 56' 26.9''\text{E}$). They formed $\sim 250\text{ m}$ from the cave entrance and $\sim 20\text{ m}$ from each other. Sample BA04 originates from Bukit Assam ($4^\circ 15' 18''\text{N}$, $114^\circ 57' 34''\text{E}$), which lies $\sim 6\text{ km}$ NNE from Snail Shell Cave. BA04 formed $>500\text{ m}$ from the cave entrance.

32. Edwards, R. L., Chen, J. H., Ku, T. L. & Wasserburg, G. J. Precise timing of the last interglacial period from mass-spectrometric determination of Th-230 in corals. *Science* **236**, 1547–1553 (1987).
33. Richards, D. A. & Dorale, J. A. in *Uranium-Series Geochemistry* (eds Bourdon, B., Henderson, G. M., Lundstrom, C. C. & Turner, S. P.) 407–460 (Reviews in Mineralogy & Geochemistry Vol. 52, Mineralogical Society of America, Chantilly, Virginia, 2003).
34. Robinson, L. F., Henderson, G. M. & Slowey, N. C. U-Th dating of marine isotope stage 7 in Bahamas slope sediments. *Earth Planet. Sci. Lett.* **196**, 175–187 (2002).
35. Edwards, R. L., Chen, J. H. & Wasserburg, G. J. ^{238}U – ^{234}U – ^{230}Th – ^{232}Th systematics and the precise measurement of time over the past 500,000 years. *Earth Planet. Sci. Lett.* **81**, 175–192 (1987).

CORRIGENDUM

doi:10.1038/nature08125

Millennial-scale trends in west Pacific warm pool hydrology since the Last Glacial MaximumJudson W. Partin, Kim M. Cobb, Jess F. Adkins, Brian Clark
& Diego P. Fernandez*Nature* 449, 452–455 (2007)

A recalibration of the Caltech ^{236}U and ^{229}Th spike has determined that the original spike value used for the stalagmite chronologies yielded U–Th ages that were too old by a margin of $\sim 2\text{--}3\%$. U–Th ages and subsequent age models have been recalculated using the new spike value for all stalagmites. The new age models alter the absolute timing of certain events, but do not alter any major conclusions of the original manuscript. The highest $\delta^{18}\text{O}$ values (inferred driest conditions) now occur at 16.0 ± 0.3 kyr ago. A late deglacial $\delta^{18}\text{O}$ plateau in the Borneo stalagmite records is now centred at 13.0 ± 0.2 kyr ago. The lowest $\delta^{18}\text{O}$ values (inferred wettest conditions) now occur at 5.0 ± 0.1 kyr ago. Examples of original and adjusted dates, respectively, at 5 kyr intervals are as follows: 5280 versus 5141 (SSC01); 10581 versus 10293 (SSC01); 15673 versus 15231 (BA04); 20937 versus 20322 (BA04); and 26387 versus 25602 (SCH02). Revised chronologies are available at: <ftp://ftp.ncdc.noaa.gov/pub/data/paleo/speleothem/pacific/gunung-buda2007.xls>

LETTERS

Metal saturation in the upper mantle

Arno Rohrbach^{1,2}, Chris Ballhaus¹, Ute Golla-Schindler², Peter Ulmer³, Vadim S. Kamenetsky⁴ & Dmitry V. Kuzmin^{5,6}

The oxygen fugacity f_{O_2} of the Earth's mantle is one of the fundamental variables in mantle petrology. Through ferric–ferrous iron and carbon–hydrogen–oxygen equilibria, f_{O_2} influences the pressure–temperature positions of mantle solidi and compositions of small-degree mantle melts^{1–3}. Among other parameters, f_{O_2} affects the water storage capacity and rheology of the mantle^{4,5}. The uppermost mantle, as represented by samples and partial melts, is sufficiently oxidized to sustain volatiles, such as H₂O and CO₂, as well as carbonatitic melts^{6,7}, but it is not known whether the shallow mantle is representative of the entire upper mantle. Using high-pressure experiments, we show here that large parts of the asthenosphere are likely to be metal-saturated. We found that pyroxene and garnet synthesized at >7 GPa in equilibrium with metallic Fe can incorporate sufficient ferric iron that the mantle at >250 km depth is so reduced that an (Fe,Ni)-metal phase may be stable. Our results indicate that the oxidized nature of the upper mantle can no longer be regarded as being representative for the Earth's upper mantle as a whole and instead that oxidation is a shallow phenomenon restricted to an upper veneer only about 250 km in thickness.

Although at the time of core melt segregation the silicate Earth must have been highly reduced and in equilibrium with metallic iron, the Earth's upper mantle is now oxidized. Relative to the iron–wüstite reference, oxygen fugacities (f_{O_2}) recorded by upper-mantle rocks and mantle-derived melts range from 3 to 6 log units above the iron–wüstite equilibrium⁸. Equilibrium with (Fe,Ni) metal at the time of core formation would have afforded an f_{O_2} of about 2 log units below the iron–wüstite equilibrium⁹. Hence, shallow upper mantle seems to have experienced oxidation by 5 to 8 orders of magnitude in f_{O_2} .

In the shallow mantle, f_{O_2} is monitored by ferric–ferrous iron equilibria such as $6Fe_2SiO_4$ (in olivine) + $O_2 = 2Fe_3O_4$ (in spinel) + $3Fe_2Si_2O_6$ (in pyroxene) or $4Fe_2SiO_4$ (in olivine) + $Fe_2Si_2O_6$ (in pyroxene) + $O_2 = 2Fe^{2+}_3Fe^{3+}_2Si_3O_{12}$ (in garnet)^{10–12}. f_{O_2} is dependent on bulk composition; the higher the Fe_2O_3/FeO bulk ratio, the more oxidized the respective mantle region will be. f_{O_2} is also pressure–temperature dependent. If at a given bulk Fe_2O_3/FeO , increasing pressure stabilizes phases that fractionate ferric iron (such as spinel or garnet), then the pressure will lower the activities of the Fe^{3+} components, causing reduction and superimposing on bulk compositional f_{O_2} effects a systematic, depth-related change in f_{O_2} . The general f_{O_2} –depth trend in the mantle is believed to be towards reduction^{13,14}.

To establish a redox profile through the upper mantle, we have equilibrated a model mantle composition (see Supplementary Information) in Fe–metal capsules to 14 GPa, corresponding to a depth of about 450 km. The starting composition was depleted relative to primitive mantle¹⁵ by 30% in normative olivine and enriched in FeO to a molar Mg/(Mg + Fe) bulk ratio of 0.5, to stabilize

ferric–iron-fractionating phases like pyroxene and garnet and to raise the iron detection limit for electron energy loss spectroscopy (EELS) analysis. Before experimentation, the starting composition was sintered at 1,150 °C in CO–CO₂ atmosphere at an f_{O_2} near iron–wüstite, to render it free of ferric iron. All experiments were performed in Fe–metal capsules from 1 to 14 GPa and 1,220 to 1,650 °C. The f_{O_2} at run conditions, ranging from 0.5 to 1.3 log units below the iron–wüstite equilibrium, was deduced from the FeO contents of pyroxene and garnet in equilibrium with metallic Fe, assuming ideal ionic solution models. Run products were analysed for major elements and then thinned to electron transparency. Pyroxene, garnet and majorite solid solutions were then analysed for their $Fe^{3+}/\Sigma Fe$ ratios using EELS^{16,17}.

At 1 GPa, stable silicate phases are olivine and two pyroxenes. From 3 to 6 GPa, subcalcic pyroxene coexists with garnet. In addition, all experimental charges are peppered with micrometre-sized metallic Fe grains (Fig. 1), suggesting that redox equilibrium with metallic Fe was attained. Al^{3+} in pyroxene falls with increasing pressure and modal garnet increases according to MA_2SiO_6 (in clinopyroxene) + $M_2Si_2O_6$ (in clinopyroxene) = $M_3Al_2Si_3O_{12}$ (in garnet) where M =

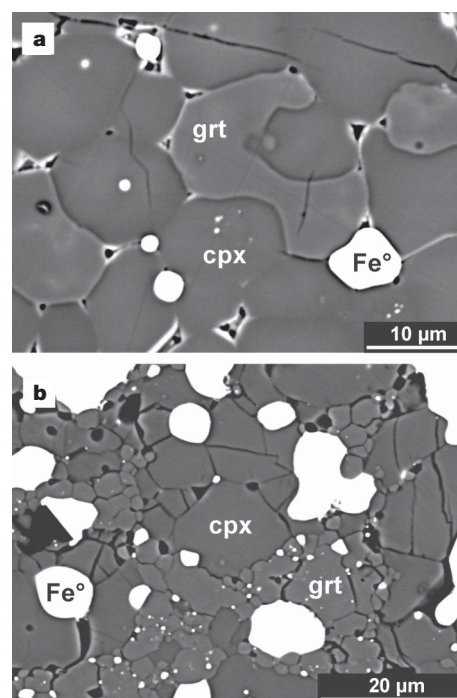


Figure 1 | Backscattered images of run products. **a**, 3 GPa experiment. **b**, 12 GPa experiment. Phases present are clinopyroxene (cpx), garnet (grt), metallic iron (Fe°) and minor amounts of partial melt.

¹Mineralogisches-Petrologisches Institut und Museum, Universität Bonn, Poppelsdorfer Schloss, 53115 Bonn, Germany. ²Institut für Mineralogie, Universität Münster, Corrensstrasse 24, 48149 Münster, Germany. ³Institut für Mineralogie und Petrographie, ETH Zürich, Clausiusstrasse 25, 8092 Zürich, Switzerland. ⁴ARC Centre of Excellence in Ore Deposits and School of Earth Sciences, University of Tasmania, Hobart, Tasmania 7001, Australia. ⁵Max-Planck-Institut für Chemie, Abt. Kosmochemie, 55128 Mainz, Germany. ⁶Institute of Geology and Mineralogy SB RAS, Novosibirsk 630090, Russia.

Mg^{2+} , Fe^{2+} and Ca^{2+} (Fig. 2a). At 7 GPa, pyroxene starts dissolving in garnet according to $2x\text{M}_2\text{Si}_2\text{O}_6$ (in clinopyroxene) + $\text{M}_3\text{Al}_2\text{Si}_3\text{O}_{12}$ (in garnet) = $[\text{M}_3\text{Al}_2\text{Si}_3\text{O}_{12} \cdot x\text{M}_3(\text{MSi})\text{Si}_3\text{O}_{12}]$ (majorite_{solid-solution}) until in the highest-pressure run (14 GPa), majorite is the only crystalline silicate. Interestingly, the onset of majorite substitution is independent of bulk $\text{Mg}/(\text{Mg}+\text{Fe})$ and occurs at the same pressure as in the more magnesian bulk composition of ref. 18 (Fig. 2b).

Figure 3 shows pressure-dependent changes in $\text{Fe}^{3+}/\Sigma\text{Fe}$ in subcalcic pyroxene and garnet. Below 6 GPa, $\text{Fe}^{3+}/\Sigma\text{Fe}$ contents are pressure-insensitive but above 7 GPa, $\text{Fe}^{3+}/\Sigma\text{Fe}$ increase rapidly, up to 0.34 at 14 GPa. Generally, garnet has higher $\text{Fe}^{3+}/\Sigma\text{Fe}$ ratios than pyroxene, as in many natural garnet peridotites¹⁹. We note a stringent correlation of $\text{Fe}^{3+}/\Sigma\text{Fe}$ with majorite component in garnet (Fig. 4), expressed as a Si_{3+x} excess over stoichiometric garnet (for which the number of Si atoms is 3). Apparently, the ability of garnet to fractionate Fe^{3+} increases with majorite substitution.

We suggest that not only the lower mantle and the transition zone^{13,14,20–23}, but also the lower half of the upper mantle is metal-saturated. Fe-metal saturation will set in when the mantle silicates in equilibrium with metallic Fe can fractionate more Fe_2O_3 than is present in the fertile upper mantle. We can approximate the depth at which this is likely to happen. At 8 GPa, fertile mantle with 4.5 wt% Al_2O_3 and 3.7 wt% CaO (ref. 15) will crystallize about 20 wt% majoritic garnet, 15 wt% subcalcic clinopyroxene, and 65 wt% olivine (assuming all Al_2O_3 fractionates into garnet and all CaO into pyroxene). A typical iron content in garnet from garnet peridotite, calculated as FeO, is 8 to 10 wt% (refs 10, 19). In our 8 GPa garnets, 15 mol% of total Fe is ferric iron. Therefore, 20 wt% majoritic garnet with an average 9 wt% FeO (refs 10, 19) may fractionate about 2,400 p.p.m. Fe_2O_3 , and this is in equilibrium with metallic Fe. Fertile upper mantle at shallow pressure contains about 2,000 p.p.m. Fe_2O_3 and about 8 wt% FeO (ref. 15). Hence, that same composition compressed to 7 to 8 GPa will be Fe-metal-saturated.

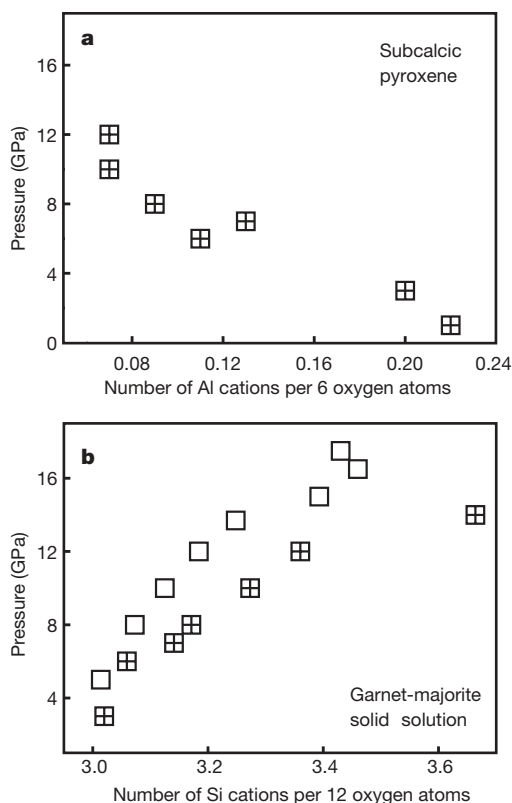


Figure 2 | Experimental clinopyroxene (a) and garnet-majorite solid solution (b) compositions. Crossed symbols, this study (bulk $\text{Mg}/(\text{Mg}+\text{Fe}) = 0.5$); open symbols, ref. 18 with bulk $\text{Mg}/(\text{Mg}+\text{Fe}) = 0.9$. Onset of majorite substitution in garnet at about 7 GPa is independent of bulk $\text{Mg}/(\text{Mg}+\text{Fe})$.

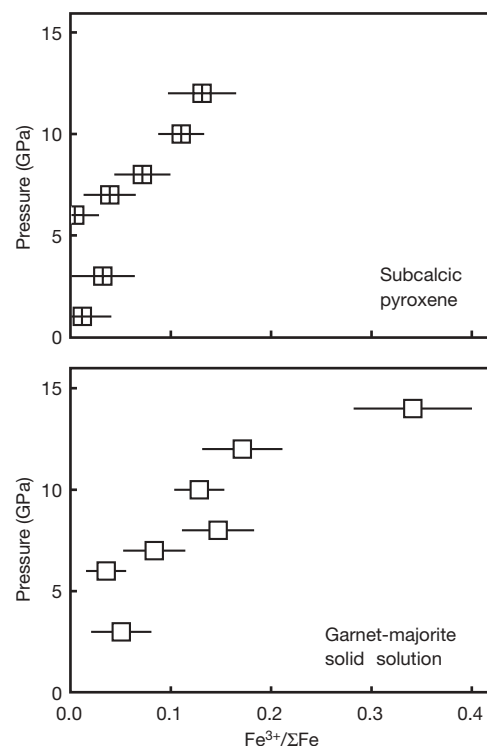


Figure 3 | Pressure effect on the $\text{Fe}^{3+}/\Sigma\text{Fe}$ atomic ratios of pyroxene and garnet (redox equilibrium with metallic Fe). Averages of 8 to 15 analyses per symbol. Error bars are standard errors of the mean with 95% confidence interval and include variations among analyses and uncertainties of the universal curve parameters of ref. 16.

Implicit in this calculation is that our FeO-enriched, olivine-depleted model composition correctly represents reactions among natural mantle phases. For example, if with increasing bulk FeO contents molar $\text{Fe}^{3+}/\Sigma\text{Fe}$ increased, as noted in ref. 24, we would have to correct our calculated level of metal saturation to greater depths because natural garnets are more magnesian. This does not seem to be the case, however. The garnets from the 8 GPa run and the 14 GPa run are the most magnesian (owing to some silicate melt lost from the charges), and yet they are within the $\text{Fe}^{3+}/\Sigma\text{Fe}$ –pressure trend in Fig. 3b or even enriched in ferric iron (14 GPa). Also, we note that our calculation is generous in that it ignores the ferric iron fractionated by clinopyroxene (Fig. 3a), so the depth of metal saturation derived above is very realistic.

With metal saturation in the upper mantle, one could be inclined to interpret the chondritic highly siderophile element (HSE) and Os

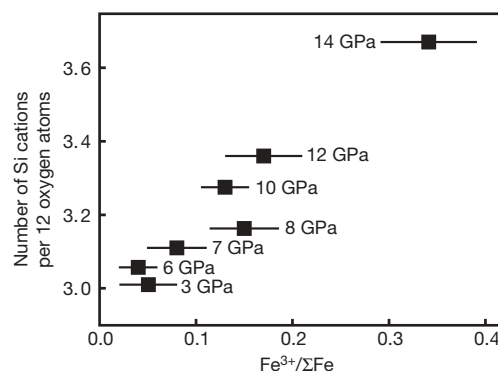


Figure 4 | Correlation between the $\text{Fe}^{3+}/\Sigma\text{Fe}$ atomic ratio and the Si_{3+x} excess in garnet-majorite solid solutions. Error bars are standard errors of the mean with 95% confidence interval and include variations among analyses and uncertainties of the universal curve parameters of ref. 16.

isotopic signatures^{15,25,26} of the mantle in terms of a “ghost signature” of a stranded core melt fraction²⁷, stable at depth but oxidized during upward convection. The oxygen source would be Fe³⁺ derived from majorite breakdown. Assuming this is true, we could use the HSE abundances of the Earth’s mantle and core¹⁵ to broadly constrain, via mass balance, the amount of an (Fe,Ni)-metal phase at depth. This mass balance gives about 1,400 p.p.m. Fe metal. If we wanted to oxidize this amount of metal to FeO by majorite breakdown, to produce an HSE “ghost signature”, we would require about 4,000 p.p.m. Fe₂O₃. This is about double the amount of Fe₂O₃ calculated above from majorite compositions. On this basis, it seems unlikely that metal saturation at depth is a relict from incomplete core formation. Also note that the upper mantle has Ni–Co overabundances²⁷, which in terms of absolute concentration, are more serious than the HSE overabundances. Basically, to account for the 2,370 p.p.m. NiO of the upper mantle by oxidation of a metal phase, in addition to the HSE overabundances, that metal would have to be nearly pure Ni, and this is quite unlike the composition of the outer core—unless the metal/silicate partition coefficients for Ni and Co decrease strongly with increasing pressure.

A metal phase at depth will influence petrologic processes in the upper mantle. In the presence of Fe-rich metal, a carbon–hydrogen–oxygen fluid will be CH₄–H₂ with negligible CO₂ and H₂O (ref. 6). Carbonates are presumably unstable²⁸ but if CH₄–H₂ fluids are decompressed they may react with Fe₂O₃ released by majorite breakdown to CO₂ and H₂O, lowering the melting temperature and inducing redox melting¹. The low-velocity region under mid-ocean ridges, which ref. 3 related to incipient CO₂ ± H₂O-triggered melting, may coincide with the depth level at which we expect a metal phase to become unstable, that is, at which carbon–hydrogen–oxygen speciations would be shifted from CH₄–H₂ to H₂O–CO₂ and induce small-degree melting. Metal saturation may also limit the amount of water to be stored in nominally anhydrous minerals because carbon–hydrogen–oxygen fluids in equilibrium with metallic Fe will be H₂O-poor⁶. Clearly, experiments are needed to test the water-storage capacity of nominally anhydrous minerals under Fe-metal-saturated conditions with CH₄–H₂ fluid.

Is there independent evidence from natural samples for a highly reduced upper mantle? In garnet peridotite xenoliths a continuous decline in relative *f*_{O₂} has been noted¹⁴ with increasing pressure to 6 GPa. Clearly, a shallow mantle trend towards reduction is encouraging for metal precipitation at greater depths. Direct evidence for metal saturation comes from Fe metal and Fe_xC carbide inclusions in diamonds^{29,30}, but one may speculate whether such inclusions record ambient *f*_{O₂} mantle conditions or local, short-lived redox perturbations when these diamonds grew²⁹. Primitive mantle melts from metal saturation depths (>250 km) seem to be rare (see also Supplementary Information). Either the necessary depths are not normally tapped or deep mantle melting itself is oxidizing, that is, triggered by oxidation¹. We note, however, that recently reported kimberlites are reduced as much as 5 log units below the nickel–NiO buffer, which is not far from (Fe,Ni)-metal saturation^{31,32}.

Received 12 March; accepted 8 August 2007.

1. Taylor, W. R. & Green, D. H. in *Magmatic Processes: Physicochemical Principles* (ed. Mysen, B. O.) 121–138 (Geochemical Society USA Special Publication 1, University Park, Pennsylvania, 1987).
2. Ballhaus, C., Berry, R. F. & Green, D. H. Oxygen fugacity controls in the Earth’s upper mantle. *Nature* **348**, 437–440 (1990).
3. Dasgupta, R. & Hirschmann, M. M. Melting in the Earth’s deep upper mantle caused by carbon dioxide. *Nature* **440**, 659–662 (2006).
4. Kohlstedt, D. L., Keppler, H. & Rubie, D. C. Solubility of water in the alpha, beta and gamma phases of (Mg,Fe)₂SiO₄. *Contrib. Mineral. Petrol.* **123**, 345–357 (1996).
5. Keppler, H. & Rauch, M. Water solubility in nominally anhydrous minerals measured by FTIR and ¹H MAS NMR; the effect of sample preparation. *Phys. Chem. Miner.* **27**, 371–376 (2000).

6. Matveev, S., Ballhaus, C., Fricke, K., Trunckenbrodt, J. & Ziegenbein, D. CHO volatiles under upper mantle conditions. I. Experimental results. *Geochim. Cosmochim. Acta* **61**, 3081–3088 (1997).
7. Wallace, M. E. & Green, D. H. An experimental determination of primary carbonatite magma composition. *Nature* **335**, 343–346 (1988).
8. Frost, B. R. in *Oxide Minerals: Petrologic and Magnetic Significance* (ed. Lindsley, D. H.) 1–9 *Reviews in Mineralogy* (ed. Ribbe, P. H.) Vol. 25 (Mineralogical Society of America, Washington DC, 1991).
9. O’Neill, H. St. C. The origin of the Moon and the early history of the Earth – A chemical model. Part 2: The Earth. *Geochim. Cosmochim. Acta* **55**, 1159–1172 (1991).
10. Luth, R. W., Virgo, D., Boyd, F. R. & Wood, B. J. Ferric iron in mantle-derived garnets. *Contrib. Mineral. Petrol.* **104**, 56–72 (1990).
11. Ballhaus, C., Berry, R. F. & Green, D. H. Experimental calibration of the olivine–orthopyroxene–spinel oxygen barometer—implications for oxygen fugacity in the Earth’s upper mantle. *Contrib. Mineral. Petrol.* **107**, 27–40 (1991).
12. Wood, B. J., Bryndzia, L. T. & Johnson, K. E. Mantle oxidation state and its relationship to tectonic environment and fluid speciation. *Science* **248**, 337–345 (1990).
13. Ballhaus, C. Is the upper mantle metal-saturated? *Earth Planet. Sci. Lett.* **132**, 75–86 (1995).
14. Woodland, A. B. & Koch, M. Variation in the oxygen fugacity with depth in the upper mantle beneath the Kaapvaal craton, Southern Africa. *Earth Planet. Sci. Lett.* **214**, 295–310 (2003).
15. Palme, H. & O’Neill, H. St. C. in *The Mantle and Core* (ed. Carlson, R. W.) 1–38 *Treatise on Geochemistry* (eds Holland, H. D. & Turekian K. K.) Vol. 2 (Elsevier–Pergamon, Oxford, 2003).
16. van Aken, P. A. & Liebscher, B. Quantification of ferrous/ferric ratios in minerals: new evaluation schemes of Fe L₂₃ electron energy-loss near-edge spectra. *Phys. Chem. Miner.* **29**, 188–200 (2002).
17. van Aken, P. A., Liebscher, B. & Styrsky, V. J. Quantitative determination of iron oxidation states in minerals using Fe L₂₃-edge electron energy-loss near-edge structure spectroscopy. *Phys. Chem. Miner.* **25**, 323–327 (1998).
18. Irifune, T. An experimental investigation of the pyroxene–garnet transformation in a pyrolyte composition and its bearing on the constitution of the mantle. *Phys. Earth Planet. Inter.* **45**, 324–336 (1987).
19. Canil, D. & O’Neill, H. St. C. Distribution of ferric iron in some upper-mantle assemblages. *J. Petrol.* **37**, 609–635 (1996).
20. O’Neill, H. St. C. et al. Mössbauer spectroscopy of mantle transition zone phases and determination of minimum Fe³⁺ content. *Am. Mineral.* **78**, 456–460 (1993).
21. Frost, D. J. et al. Experimental evidence for the existence of iron-rich metal in the Earth’s lower mantle. *Nature* **428**, 409–412 (2004).
22. Wade, J. & Wood, B. J. Core formation and the oxidation state of the Earth. *Earth Planet. Sci. Lett.* **236**, 78–95 (2005).
23. Sinmyo, R., Hirose, K., O’Neill, H. St. C. & Okunishi, E. Ferric iron in Al-bearing post-perovskite. *Geophys. Res. Lett.* **33**, L12513, doi:10.1029/2006GL025858 (2006).
24. McCammon, C. A. & Ross, N. L. Crystal chemistry of ferric iron in (Mg,Fe)(Si,Al)O₃ majorite with implications for the transition zone. *Phys. Chem. Miner.* **30**, 206–216 (2003).
25. Ringwood, A. E. Origin of chondrites. *Nature* **207**, 701–704 (1965).
26. Meisel, T., Walker, R. J. & Morgan, J. W. The osmium isotopic composition of the Earth’s primitive upper mantle. *Nature* **383**, 517–520 (1996).
27. Jones, J. H. & Drake, M. J. Geochemical constraints on core formation in the Earth. *Nature* **322**, 221–228 (1986).
28. Eggler, D. H. & Baker, D. R. in *High Pressure Research in Geophysics* (eds Akimoto, S. & Manghnani, M. H.) 237–250 (Center Academic, Tokyo, 1982).
29. Jacobs, D. E., Kronz, A. & Viljoen, K. S. Cohenite, native iron and troilite inclusions in garnets from polycrystalline diamond aggregates. *Contrib. Mineral. Petrol.* **146**, 566–576 (2004).
30. Stachel, T., Harris, J. W. & Brey, G. P. Rare and unusual mineral inclusions in diamonds from Mwadui, Tanzania. *Contrib. Mineral. Petrol.* **132**, 34–47 (1998).
31. Bellis, A. J. & Canil, D. Ferric iron in CaTiO₃ perovskite as an oxygen barometer for kimberlitic magmas. I: Experimental calibration. *J. Petrol.* **48**, 219–230 (2007).
32. Canil, D. & Bellis, A. J. Ferric iron in CaTiO₃ perovskite as an oxygen barometer for kimberlite magmas. II: Applications. *J. Petrol.* **48**, 231–252 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Museum of Natural History in London for providing Karoo picrite samples from Malawi. Comments on the manuscript by R. Frost and D. Canil and D. Frost improved the paper. Financial support by the German Research Council to C.B. is gratefully acknowledged.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.R. (rohrbaa@web.de) or C.B. (ballhaus@uni-bonn.de).

Fire and flood management of coastal swamp enabled first rice paddy cultivation in east China

Y. Zong¹, Z. Chen², J. B. Innes¹, C. Chen³, Z. Wang² & H. Wang⁴

The adoption of cereal cultivation was one of the most important cultural processes in history, marking the transition from hunting and gathering by Mesolithic foragers to the food-producing economy of Neolithic farmers¹. In the Lower Yangtze region of China, a centre of rice domestication², the timing and system of initial rice cultivation remain unclear. Here we report detailed evidence from Kuahuqiao that reveals the precise cultural and environmental context of rice cultivation at this earliest known Neolithic site in eastern China, 7,700 calibrated years before present (cal. yr BP). Pollen, algal, fungal spore and micro-charcoal data from sediments demonstrate that these Neolithic communities selected low-land swamps for their rice cultivation and settlement, using fire to clear alder-dominated wetland scrub and prepare the site for occupation, then to maintain wet grassland vegetation of paddy type. Regular flooding by slightly brackish water was probably controlled by 'bunding' to maintain crop yields. The site's exploitation ceased when it was overwhelmed by marine inundation 7,550 cal. yr BP. Our results establish that rice cultivation began in coastal wetlands of eastern China, an ecosystem vulnerable to coastal change but of high fertility and productivity, attractions maximized for about two centuries by sustained high levels of cultural management of the environment.

The major climate amelioration from the late Pleistocene to the Holocene epoch markedly altered global ecological systems, prompting changes in human environmental adaptations and, in favourable regions including China^{3,4}, enabling transitions from hunting and gathering to domestication. Mesolithic foragers gathered aquatic perennial wild rice (*Oryza rufipogon*) in the middle Yangtze basin⁵ and the lower Yangtze region⁶ from the beginning of the Holocene epoch, an adaptation resulting from its availability due to the favourable environmental conditions of a warm, humid, monsoonal climate⁷. Such exploitation of wild rice was a prelude to its cultivation and then domestication⁸ in these two core areas⁹. Understanding the process leading to rice domestication is constrained by ignorance of the precise ecological conditions that accompanied its initial cultivation. Our research at the wetland site of Kuahuqiao in the lower Yangtze region (Fig. 1) has answered this fundamental question regarding the origin of wet rice cultivation systems.

Excavation of cultural layers in organic deposits at Kuahuqiao¹⁰ recovered rich assemblages of biological remains and archaeological artefacts, including a dugout canoe¹¹ and both forager- and cultivator-type bone, bamboo and wooden tools, well preserved because of their waterlogged condition and dated between 7,700 and 7,550 cal. yr BP (Table 1). Rice-tempered ceramics were common. The settlement's wooden pile dwellings prove that it was situated within wet marshes, and its excellent preservation of biological remains allows reconstruction of its economy and environment. Oysters provide evidence of the exploitation of coastal resources. Common animal bones include

domestic pig and dog but mainly comprise wild species, as do the abundant plant remains, which include many edible types, such as acorns. Hunting and gathering dominated the economy. Many grains and phytoliths of rice occur, mostly wild varieties but including large numbers of morphologically advanced (but not yet domesticated) forms¹². The cultural sediments represent the *in situ* location of first rice cultivation by hunter-gatherers establishing sedentary villages and undergoing Neolithic economic transition. Cultural activities continued for almost two centuries before the site was abandoned after inundation by mid-Holocene sea-level rise¹³ and sealed beneath marine clay.

Lakeside and coastal swamps are good candidates for the ecological zone where earliest rice cultivation occurred¹⁴, and Kuahuqiao's waterlogged sediments and cultural soils are ideal for testing this proposal, as they contain evidence of wild and early cultivated rice

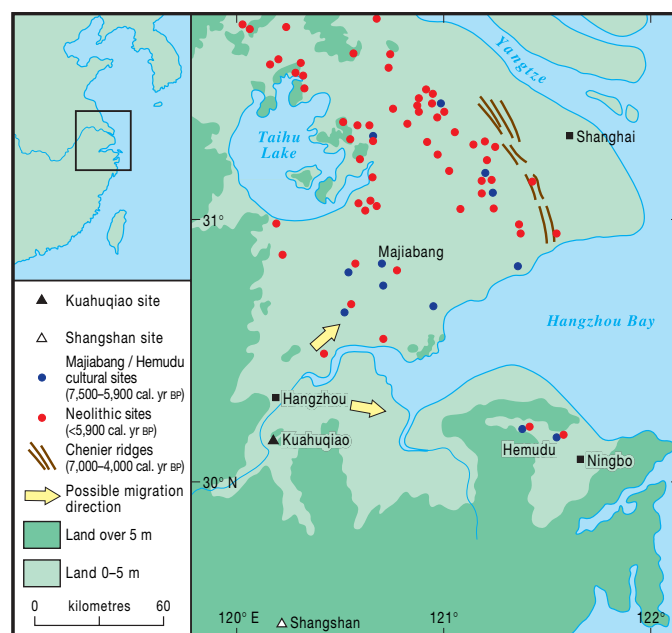


Figure 1 | Location of the Kuahuqiao site, and others, in the lower Yangtze region. The Kuahuqiao site lies at the gateway between sites in the upland valleys, south and west of Kuahuqiao, occupied by late Mesolithic and early Neolithic hunters and foragers including occupants at the Shangshan site²⁷, and sites on the coastal wetlands of Hangzhou Bay, east and north of Kuahuqiao, occupied by early Neolithic farmers such as those in Hemudu²⁵ and Majiabang⁹ (blue dots). Red dots indicate locations of Songze and Liangzhu cultural sites¹⁹. The yellow arrows indicate the possible migration directions of the early Neolithic communities. Ages shown are in calibrated years BP.

¹Geography Department, Durham University, Durham DH1 3LE, UK. ²Institute for Estuarine and Coastal Research, East China Normal University, Shanghai 200062, China.

³Department of Cultural Heritage, Fudan University, Shanghai 200433, China. ⁴Department of Geography, East China Normal University, Shanghai 200062, China.

Table 1 | Chronology and the radiocarbon dates for the Kuahuqiao site

Depth (cm below sea level)	Laboratory code	Methods	¹⁴ C age (yr BP, ± 1σ)	Material	Calibrated age (yr BP, 1σ)	Central calibrated age (yr BP, ± 1σ)	Events
195.0	From ref. 10	Radiometric	6,330 ± 190	Bulk organic	7,144–7,428	7,286 ± 142	Brackish water conditions
200.0	BA05766	AMS	6,710 ± 40	Organic fragments	7,564–7,612	7,588 ± 24	End of human activities
208.0	BA05767	AMS	6,805 ± 35	Organic fragments	7,616–7,670	7,643 ± 27	Rise in cultural NPMs
210.5	GZ1311	AMS	6,743 ± 36	Organic fragments	7,574–7,622	7,598 ± 24	Decline in reed-swamp NPMs
213.5	GZ1312	AMS	6,710 ± 31	Organic fragments	7,566–7,608	7,587 ± 21	Rise in <i>Typha</i>
224.5	GZ1314	AMS	6,752 ± 33	Charcoal	7,579–7,622	7,601 ± 22	End of large-scale burning
229.5	GZ1315	AMS	6,783 ± 32	Leaf	7,607–7,662	7,635 ± 28	Increase in reed-swamp NPMs
238.0	GZ1316	AMS	6,851 ± 33	Organic fragments	7,653–7,709	7,681 ± 28	Start of human activities. Burning and end of alder carr
238.0	BA05768	AMS	6,870 ± 40	Organic fragments	7,663–7,746	7,705 ± 42	
242.5	GZ1317	AMS	6,996 ± 33	Organic fragments	7,792–7,863	7,828 ± 36	Rise of alder carr
282.0	From ref. 10	Radiometric	8,125 ± 250	Bulk organic	8,722–9,319	9,021 ± 299	Change to freshwater conditions

AMS, accelerator mass spectrometry.

in secure archaeological contexts, while wild rice occurred naturally in the area¹⁵. Whereas previous studies of the origins of rice cultivation in China have relied on rice phytoliths^{8,16}, and only rarely using pollen¹⁷, we have used high-resolution pollen and several other types of non-pollen microfossil (NPM) analyses to provide detailed, diagnostic palaeoecological data from Kuahuqiao, summarized in Fig. 2. Eleven radiocarbon dates, concentrated in the cultural layers, provide a secure chronology for the sequence of environmental changes.

High frequencies of marine and brackish water diatoms and estuarine NPMs—mainly intertidal fungal spores and dinoflagellate

cysts—in the clays of units A and G reflect marine sedimentation. Salt-marsh herb pollen occurs in these marine units, but much of the pollen assemblage comes from freshwater communities, indicating estuarine rather than open coast environments. The radiocarbon dates indicate about 1,500 yr of mainly organic freshwater sedimentation in phases B to F, when the rate of sea-level rise was greatly reduced¹⁸, between two marine episodes. The combined pollen curve for deciduous (*Quercus*) and evergreen (*Cyclobalanopsis*) oaks represents the dominant upland vegetation (that is, sub-tropical woodland), indicating a warm, wet monsoonal climate¹⁹.

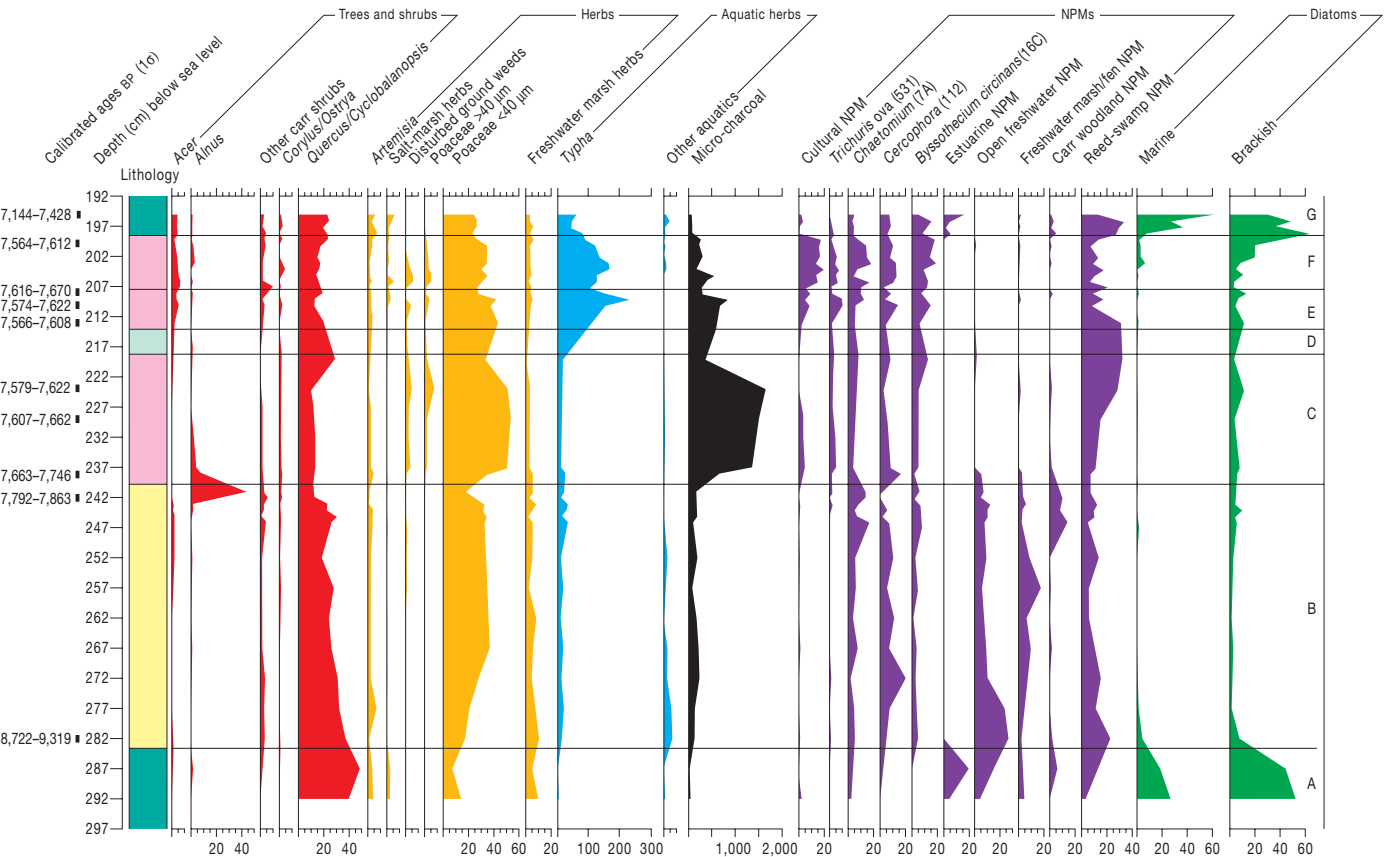


Figure 2 | Integrated palaeoecological data from Kuahuqiao. Generalized stratigraphical units and calibrated radiocarbon age ranges BP are shown. Microfossil data are subdivided into seven phases. Pollen and microscopic charcoal frequencies are calculated as percentages of a total land pollen sum including trees, shrubs and terrestrial herbs, but not aquatic herbs or pteridophyte and bryophyte spores. NPMs are shown as percentages of their

own total counted sums. Numbers after NPM taxa names refer to the standard catalogue²¹. Some curves are aggregated percentages of several taxa of similar ecological affinities. Major increases in micro-charcoal, Poaceae/reed-swamp microfossils and cultural indicators occur in phases C, E and F.

The organic clay microfossil evidence in unit B indicates fresh-water conditions, with no saltwater influence. Swamps with open-water pools are indicated by many aquatic fungi, macrophytes and algae, giving way to marsh, fen and reed-swamp wetland types. In the upper part of unit B, increases in grass (*Poaceae* <40 µm) and fresh-water marsh herb pollen indicate progressively shallower water as reed-swamp expanded. Abundant *Cercophora*, a fungal decomposer of herbaceous material, is further proof of the fen-marshland vegetation. By the end of phase B, successions culminated in the establishment of dense wetland scrub: first birch–willow (*Betula*–*Salix*) then alder (*Alnus*). Diatom evidence from here onwards shows a very slight rise in brackish water influence.

This scrub was cleared by fire at about 7,700 cal. yr BP at the start of the first cultural episode (phase C), as the curves for alder pollen, swamp woodland NPMs and aquatic/fen taxa collapse and micro-charcoal frequencies increase tenfold. Curves for cereal (*Oryza*) type pollen (thick annulus *Poaceae* >40 µm), disturbed ground weeds, 'cultural' NPMs and the ova of the parasitic worm *Trichuris*, hosted by pigs and people²⁰, begin at this time. After scrub clearance, human occupation and rice cultivation began on wet organic soils in grass–reed-swamp environments. This first intensive culture phase lasted over a century before interruption by a flooding episode and silt-clay deposition, with some woodland regeneration and reduced microfossil evidence of cultural activity (phase D). From the end of unit B throughout cultural phases C, E and F increased levels of salt-marsh herb pollen and salt-tolerant diatoms indicate regular flooding by slightly brackish water (freshwater diatoms still >85% of total). The cultural phases E and F saw intensive exploitation but reduced burning. 'Cultural' microfossil levels are very high during these two phases, including fungal spores indicative of soil disturbance and erosion, and some (*Chaetomium* and *Cercophora*) that are indicative of decomposing dung²¹. Seasonal high-tide flooding of coastal rice fields²² combined with animal grazing²³ and low-intensity controlled burning of the reed-swamp promoted the growth of cattail (*Typha*). The micro-charcoal data in the cultural phases plus the dung-associated fungi and the site's abundant pig bones suggest human management, with burning and manuring to enhance rice cultivation yields. This encouraged abundant *Typha* growth around the cultivated area, which was perhaps deliberate as *Typha* is itself an important crop for food and materials²⁴. This intensive human manipulation of the environment continued until 7,550 cal. yr BP, when marine inundation ended settlement and cultivation.

The cultural evidence shows that the forager/cultivator site at Kuahuqiao was the forerunner of a concentration of early Neolithic sites in the coastal lowlands around Taihu Lake and Hangzhou Bay (Fig. 1), including classical sites of the Chinese Neolithic, such as Hemudu^{25,26}. This area was a major centre of rice exploitation from the very early Holocene, as shown by rice husks and phytoliths in pottery fabric before 9,000 cal. yr BP at Shangshan²⁷, with the establishment at 6,500 cal. yr BP of agricultural villages, as at Hemudu, and the intensive farming of fully domesticated rice varieties during the Songze and Liangzhu Neolithic cultures after 6,000 cal. yr BP¹⁹. There is strong debate over how agriculturally meaningful the earliest Holocene rice records are^{27,28}, but rice grain morphology suggests that records such as Shangshan are of exploited wild rice varieties, with dedicated cultivation (of evolving but still pre-domestication forms¹²) starting in the cultural occupation at Kuahuqiao and then at Hemudu²⁸. The precise ecological data from Kuahuqiao are therefore critical to understanding the early stages of the forager–cultivator transition. A similar lengthy transitional period, with increasingly intensive cultivation of pre-domesticated rice varieties, also occurred in the central Yangtze valley around Lakes Poyang and Dongting at sites like Bashidang and Pengtoushan²⁹. These sites have abundant rice grain and rice-tempered ceramic evidence comparable to Kuahuqiao, whereas similar sites in that area, like Diaotonghuan, may well be forager/cultivator correlatives of Kuahuqiao. Although their precise ecological history remains

unknown, these sites provide a context for Kuahuqiao, and it is likely that the transition to dedicated rice cultivation was encouraged along the Yangtze valley by the major climate switch to benign, warm, wet conditions under the strengthening summer monsoon after 9,000 cal. yr BP³⁰, a critical environmental prompt to cultural change, permitting rice cultivation at this latitude.

The environmental evidence from Kuahuqiao not only establishes the presence of extremely early dedicated rice cultivation but allows an assessment of the earliest Neolithic rice cultivation system itself. Our detailed ecological analyses at Kuahuqiao establish that in eastern China the earliest significant rice cultivation system was located in slightly brackish coastal reed-swamps, and that, even at this early stage, rice cultivation involved very high-intensity clearance and management of the coastal marsh vegetation by fire. The abundant *Typha* stands that formed at the site, encouraged by human management activity, provided another highly productive crop, and rice cultivation may have been only one aspect of the exploitation of coastal resources that encouraged settlement, environmental manipulation and early cultivation in locations like Kuahuqiao. It is likely that floodwater input to the cultivated areas was also controlled by humans, as the proportion of tidal brackish water influence is maintained at a consistently low level throughout the later cultural phases. The earliest system of rice cultivation in China may well have been a form of 'receding-flood' water regulation, with artificial bunding used to retain some nutrient-rich seasonal floodwater, prevent major inundation and provide rice with the consistent water regime it requires. The eventual overwhelming of the site by marine inundation reflects the breakdown of this system of flood regulation and highlights the vulnerability of early rice cultivation in this productive but environmentally unstable coastal ecosystem. Our conclusion that incipient Neolithic groups in China used fire management to modify coastal wetlands for intensive early rice cultivation may pertain to other areas, such as around the lakes of the central Yangtze valley, and requires testing there.

METHODS SUMMARY

Field and laboratory methodology was designed to produce high-resolution palaeo-ecological data. Sediments for analysis were collected from an exposed section at the edge of the archaeological site, close enough to preserve evidence regarding on-site and near-site activities but far enough away from the main occupation area to avoid disturbance of the deposits' stratigraphy in antiquity. The altitude of the sampled profile was recorded relative to modern Yellow Sea Datum so that the site's relationship with past sea level could be established. Close interval sampling of the deposits was undertaken so that short timescale events could be recognised and studied. Multiple accelerator mass spectrometry (AMS) dates provide a robust and detailed chronology for the sedimentary sequence. Laboratory chemical preparations removed all mineral and organic material except for microfossils and microscopic charcoal fragments. Pollen grains were selected as the main source of evidence because they derive from a spatial variety of source areas, so that both on- and off-site vegetation can be reconstructed. Fungal spores, diatoms, other algae and invertebrate remains were also recorded, as these provide evidence sensitive to changes in local environmental conditions, including tidal brackish water influence. Micro-charcoal particles were recorded so that local fire history could be studied.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 April; accepted 31 July 2007.

1. Lu, T. L. D. *The Transition from Foraging to Farming and the Origin of Agriculture in China* (British Archaeological Reports International Series 774, John & Erica Hedges, Oxford, 1999).
2. Londo, J. P., Chiang, Y.-C., Hung, K.-H., Chiang, T.-Y. & Schaal, B. A. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl Acad. Sci. USA* **103**, 9578–9583 (2006).
3. Zhao, Z. J. & Piperno, D. R. Late Pleistocene/Holocene environments in the middle Yangtze River valley, China and rice (*Oryza sativa* L.) domestication: the phytolith evidence. *Geoarchaeology* **15**, 203–222 (2000).
4. Lu, H. *et al.* Rice domestication and climatic change: phytolith evidence from East China. *Boreas* **31**, 378–385 (2002).

5. Higham, C. F. W. & Lu, T. The origins and dispersal of rice cultivation. *Antiquity* **72**, 867–877 (1998).
 6. Zhao, Z., Pearsall, D. M. & Jiang, Q. H. in *Origins of Rice Agriculture* (eds MacNeish, R. S. & Libby, J. G.) 47–52 (Univ. Texas Press, El Paso, 1995).
 7. Wang, Y. *et al.* The Holocene Asian monsoon: links to solar changes and North Atlantic climate. *Science* **308**, 854–857 (2005).
 8. Zhao, Z. The middle Yangtze region in China is one place where rice was domesticated: phytolith evidence from the Diaotonghuan cave, northern Jiangxi. *Antiquity* **72**, 885–897 (1998).
 9. Crawford, G. W. & Shen, C. The origins of rice agriculture: recent progress in East Asia. *Antiquity* **72**, 858–866 (1998).
 10. Jiang, L. *Kuahuqiao* (Wenwu Press, Beijing, 2004).
 11. Jiang, L. & Liu, L. The discovery of an 8000-year-old dugout canoe at Kuahuqiao in the Lower Yangtze River, China. *Antiquity* **79**, Project Gallery 305 (<http://antiquity.ac.uk/projgall/liu/index.html>) (2005).
 12. Zheng, Y., Jiang, L.-P. & Zheng, J.-M. Study on the remains of ancient rice from Kuahuqiao site in Zhejiang province. *Chin. J. Rice Sci.* **18**, 119–124 (2004).
 13. Zong, Y. Mid-Holocene sea-level highstand along the southeast coast of China. *Quat. Int.* **117**, 55–67 (2004).
 14. Glover, I. C. & Higham, C. F. W. in *The Origin and Spread of Agriculture and Pastoralism in Eurasia* (ed. Harris, D. R.) 413–441 (UCL Press, London, 1996).
 15. Tang, S., Sato, Y. I. & Yu, W. Discovery of wild rice grains (*O. rufipogon*) in the Hemudu ancient carbonized rice. *Agric. Archaeol.* **3**, 88–122 (1994).
 16. Jiang, Q. Searching for evidence of early rice agriculture at prehistoric sites in China through phytolith analysis: an example from central China. *Rev. Palaeobot. Palynol.* **89**, 481–485 (1995).
 17. Huang, F. & Zhang, M. Pollen and phytolith evidence for rice cultivation during the Neolithic at Longqiuzhuang, eastern Jianghuai, China. *Veg. Hist. Archaeobot.* **9**, 161–168 (2000).
 18. Bird, M. I. *et al.* An inflection in the rate of early mid-Holocene eustatic sea-level rise: a new sea-level curve from Singapore. *Estuar. Coast. Shelf Sci.* **71**, 523–536 (2007).
 19. Chen, Z., Wang, Z., Schneiderman, J., Tao, J. & Cai, Y. Holocene climate fluctuations in the Yangtze delta of eastern China and the Neolithic response. *Holocene* **15**, 917–926 (2005).
 20. Pike, A. W. Recovery of helminth eggs from archaeological excavations, and their possible usefulness in providing evidence for the purpose of an occupation. *Nature* **219**, 303–304 (1968).
 21. van Geel, B. & Aptroot, A. Fossil ascomycetes in Quaternary deposits. *Nova Hedwigia* **82**, 313–320 (2006).
 22. Mesleard, F. J., Lepart, J., Grillas, P. & Mauchamp, A. Effects of seasonal flooding and grazing on the vegetation of former rice fields in the Rhone delta (Southern France). *Plant Ecol.* **145**, 101–114 (1999).
 23. Graf, M.-T. & Chmura, G. L. Development of modern analogues for natural, mowed and grazed grasslands using pollen assemblages and coprophilous fungi. *Rev. Palaeobot. Palynol.* **141**, 139–149 (2006).
 24. Morton, J. F. Cattails (*Typha* spp.)—weed problem or potential crop? *Econ. Bot.* **29**, 7–29 (1975).
 25. Jiang, L. *Hemudu—A Neolithic Site and its Archaeological Excavations* (Wenwu Press, Beijing, 2003).
 26. Zhao, S. & Wu, W.-T. Early Neolithic Hemudu culture along the Hangzhou estuary and the origin of domestic paddy rice in China. *Asian Perspect.* **27**, 29–34 (1988).
 27. Jiang, L. & Liu, L. New evidence for the origins of sedentism and rice domestication in the Lower Yangtze River, China. *Antiquity* **80**, 355–361 (2006).
 28. Fuller, D. Q., Harvey, E. & Qin, L. Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity* **81**, 316–331 (2007).
 29. Pei, A. Notes on new advancements and revelations in the agricultural archaeology of early rice domestication in the Dongting Lake region. *Antiquity* **72**, 878–885 (1998).
 30. Jarvis, D. Pollen evidence of changing Holocene monsoon climate in Sichuan province, China. *Quat. Res.* **39**, 325–337 (1993).
- Acknowledgements** This research is funded by the Leverhulme Trust (to Y.Z.) and assisted by a visiting professorship by the East China Normal University, Shanghai, China (to Y.Z.).
- Author Contributions** Y.Z., Z.C. and C.C. worked on the project design. Z.C., C.C., Z.W. and H.W. helped with fieldwork and sample collection. Z.C. provided the radiocarbon dates. J.B.I. and Y.Z. analysed the microfossil data. Y.Z., Z.C. and J.B.I. wrote the paper. All authors discussed the results and commented on the manuscript.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Y.Z. (y.q.zong@durham.ac.uk) or J.B.I. (j.b.innes@durham.ac.uk).

METHODS

Further details of the Kuahuqiao site. The site is located about 5 km south of Hangzhou City, between the hilly area to the south and west and the coastal lowland of Hangzhou Bay to the east and north, and is at the gateway between the catchment basin of the rivers and the coast. The archaeological site lies on the edge of a low-lying depression or a lake, with ground altitudes around 3.8 m above mean sea level, between two lines of hills.

The lacustrine sediments at the site were dug for brick making in the 1970s. As a result, part of the archaeological site was destroyed, and no cemetery was found during later excavations. The site was excavated for the first time in 1990, and excavated again in 2001 and 2002, by the Zhejiang Provincial Institute of Cultural Relics and Archaeology, an area of 1,080 m² in total¹⁰. At the centre of the site the cultural layer is about 3 m thick, covering the living quarter of the site from which at least four piled dwellings were found. The current research investigates the eastern edge of the site where the cultural layer is about 0.5 m thick and from which palaeoecological reconstruction is more reliable. The eastern profile (facing west) of trench T0512 was sampled in 2005, and this profile is about 8 m from the location where a dugout canoe was found and dated to approximately 8,000 cal. yr BP¹¹. The sedimentary sequences at the profile are recorded against the Yellow Sea Datum: altitude -97 to -161 cm, grey, laminated fine sands, silt and clay, with abundant foraminifera, interpreted as tidal flat deposits; altitude -161 to -199 cm, brownish grey, uniform silt and clay, with abundant foraminifera, interpreted as tidal flat deposits; altitude -199 to -239 cm, dark brown silt and clay with abundant organic detritus and cultural remains, cultural layers C, E and F; altitude -239 to -283 cm, blue to grey, soft clay, turning brownish towards the top, with abundant plant fragments, interpreted as lacustrine deposits; altitude -283 to -297 cm, blue to grey, silt and clay, changing into yellowish sands and gravels at the base.

Geographically, the Kuahuqiao site lies at the gateway between the hilly areas and the coastal lowlands. Archaeologically, the Kuahuqiao site marks the beginning of the cultural transition between foraging Mesolithic hunter-gatherers and Neolithic food-producing farmers. In Kuahuqiao, over 1,000 rice grains were found, some of the grains show characteristics of domestic form, dated to 7,000 to 8,000 cal. yr BP¹⁰. Thus the Kuahuqiao site is considered important as it bridges the transition between sedentary foraging in Shangshan, approximately 90 km south of Kuahuqiao, where pottery pastes were found tempered with charred plants including rice husks and dated to about 10,000 to 8,000 cal. yr BP²⁷, and full domestication in Hemudu, about 140 km east of Kuahuqiao, where a large quantity of rice grains and husks were found and dated to approximately 7,000 to 6,000 cal. yr BP⁹. Recent examination of rice grains from the Lower Yangtze region suggests that there is an extended period, a millennium or more, of pre-domestication cultivation, which exists between a much longer period of wild rice use by foragers and the full domestication around 6,000 cal. yr BP²⁸. The

cultural activities found in Kuahuqiao fall right within this transitional pre-domestication cultivation period.

According to ref. 10, other archaeological findings from Kuahuqiao include piled dwelling and storage pits. There are a large number of ceramics, mostly tempered with charred plants and rice husks. There are also a large number of bones from a variety of animals, birds and fish, as well as oyster shells and a variety of nuts and fruit seeds, acorns in particular. These remains suggest a wide range of activities of the Kuahuqiao community, including hunting, fishing and gathering as well as cultivation.

Laboratory methods. Microfossils (that is, <180 µm in size) from the sampled section were counted using a Nikon stereomicroscope at magnifications of ×400, using ×600 oil immersion lenses for identification of critical features. Thirty-six sub-samples were taken for analysis at no more than 5 cm vertical increments, closing to 1 cm intervals through most of the cultural layers. Samples were prepared for pollen analysis using standard preparation techniques including alkali digestion, hydrofluoric acid digestion and acetolysis³¹. Microfossil concentration was good although preservation was variable, and a minimum of 200 pollen grains was counted at each sampled level, after which the rest of the microscope slide was scanned for rare or indicator types and to confirm the presence or absence of cultigen pollen. Non-pollen microfossils (NPMs), mainly comprising fungal spores and remains of algae and invertebrates, were also recorded with at least 100 identified at each level and usually many more. NPMs have been proven to be accurately represented on slides prepared for pollen analysis³². Microscopic charcoal fragments were counted, their numbers at each level calculated by their aggregated size relative to a unit of measurement equal to an average-sized pollen grain, providing a pollen/micro-charcoal ratio³³. Diatom counts were also made, with taxa classified according to a halobian system based on salinity tolerances³⁴. Frequency curves for selected types of these proxy data are shown in Fig. 2, correlated with radiocarbon assays and stratigraphical units. The calibration to calendar dates used the Calibration Program Calib 5.1 (ref. 35).

31. Moore, P. D., Webb, J. & Collinson, M. *Pollen Analysis* (Blackwell, Oxford, 1991).
32. Clarke, C. M. In *Aspects of Archaeological Palynology: Methodology and Applications* (eds Davis, O. K. & Overs, J.) 53–62 (AASP Contributions Series 29, 1994).
33. Innes, J. B. & Simmons, I. G. Mid Holocene charcoal stratigraphy, fire history and palaeoecology at North Gill, North York Moors, UK. *Palaeogeog. Palaeoclimatol. Palaeoecol.* **164**, 151–165 (2000).
34. Vos, P. C. & de Wolf, H. Diatoms as a tool for reconstructing sedimentary environments in coastal wetlands: methodological aspects. *Hydrobiologia* **269/270**, 285–296 (1993).
35. Stuiver, M., Reimer, P. J. & Braziunas, T. F. High-resolution radiocarbon age calibration for terrestrial and marine samples. *Radiocarbon* **40**, 215–230 (1998).

The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla

The French–Italian Public Consortium for Grapevine Genome Characterization*

The analysis of the first plant genomes provided unexpected evidence for genome duplication events in species that had previously been considered as true diploids on the basis of their genetics^{1–3}. These polyploidization events may have had important consequences in plant evolution, in particular for species radiation and adaptation and for the modulation of functional capacities^{4–10}. Here we report a high-quality draft of the genome sequence of grapevine (*Vitis vinifera*) obtained from a highly homozygous genotype. The draft sequence of the grapevine genome is the fourth one produced so far for flowering plants, the second for a woody species and the first for a fruit crop (cultivated for both fruit and beverage). Grapevine was selected because of its important place in the cultural heritage of humanity beginning during the Neolithic period¹¹. Several large expansions of gene families with roles in aromatic features are observed. The grapevine genome has not undergone recent genome duplication, thus enabling the discovery of ancestral traits and features of the genetic organization of flowering plants. This analysis reveals the contribution of three ancestral genomes to the grapevine haploid content. This ancestral arrangement is common to many dicotyledonous plants but is absent from the genome of rice, which is a monocotyledon. Furthermore, we explain the chronology of previously described whole-genome duplication events in the evolution of flowering plants.

All grapevine varieties are highly heterozygous; preliminary data showed that there was as much as 13% sequence divergence between alleles, which would hinder reliable contig assembly when a whole-genome shotgun strategy was used for sequencing. Our consortium therefore selected the grapevine PN40024 genotype for sequencing. This line, originally derived from Pinot Noir, has been bred close to full homozygosity (estimated at about 93%) by successive selfings, permitting a high-quality whole-genome shotgun assembly.

A total of 6.2 million end-reads were produced by our consortium, representing an 8.4-fold coverage of the genome. Within the assembly, performed with Arachne¹², 316 supercontigs represent putative allelic haplotypes that constitute 11.6 million bases (Mb). These values are in good fit with the 7% residual heterozygosity of PN40024 assessed by using genetic markers. When considering only one of the haplotypes in each heterozygous region, the assembly (Table 1a) consists of 19,577 contigs (N_{50} = 65.9 kilobases (kb), where N_{50} corresponds to the size of the shorter supercontig or contig in a subset representing half of the assembly size) and 3,514 supercontigs (N_{50} = 2.07 Mb) totalling 487 Mb. This value is close to the 475 Mb previously reported for the grapevine genome size¹³.

Using a set of 409 molecular markers from the reference grapevine map¹⁴, 69% of the assembled 487 Mb, arranged into 45 ultracontigs

Table 1 | Global statistics on the genome of *Vitis vinifera*

(a) Assembly						
	Status	Number	N_{50} (kb)	Longest (kb)	Size (Mb)	Percentage of the assembly
Contigs	All	19,577	65.9	557	467.5	–
Supercontigs	All	3,514	2,065	12,675	487.1	100
	Anchored on chromosomes	191	3,189	12,675	335.6	68.9
	Anchored on chromosomes and oriented	143	3,827	12,675	296.9	60.9
(b) Annotation						
	Number	Median size (bp)	Total length (Mb)	Percentage of the genome	%GC	
Gene	30,434	3,399	225.6	46.3	36.2	
Exons CDS	149,351	130	33.6	6.9	44.5	
Introns CDS	118,917	213	178.6	36.7	34.7	
Intergenic	30,453	3,544	261.5	34.7	33.0	
tRNA*	600	73	0.04	NS	43.0	
miRNA†	164	103.5	0.002	NS	35.9	
(c) Orthology						
	Number of orthologous proteins	Mean identity (%)				
<i>P. trichocarpa</i>	12,996	72.7				
<i>A. thaliana</i>	11,404	65.5				
<i>O. sativa</i>	9,731	59.8				
Common to eudicotyledons‡	10,547					
Common to Magnoliophyta§	8,121					

* Transfer RNA (tRNA) values were computed on exons.

† Micro RNAs (miRNAs) are members of known conserved miRNA families.

‡ Eudicotyledons are represented by *P. trichocarpa* and *A. thaliana*.

§ Magnoliophyta (most flowering plants) are represented by *P. trichocarpa*, *A. thaliana* and *O. sativa*.

*A list of participants and their affiliations appears at the end of the paper.

and 51 single supercontigs, were anchored along the 19 linkage groups. Thirty-seven ultracontigs and 22 single supercontigs were oriented, representing 61% of the genome assembly (Supplementary Tables 2 and 3).

This assembly has been annotated by using a combination of evidence. The major features of the genome annotation are presented in Table 1b. The 8.4-fold draft sequence of the grapevine genome contains a set of 30,434 protein-coding genes (an average of 372 codons and 5 exons per gene). This value is considerably lower than the 45,555 protein-coding genes reported for the poplar (*Populus trichocarpa*) genome, which has a similar size, at 485 Mb (ref. 1), and even lower than the 37,544 protein-coding genes identified in the 389 Mb of the rice genome².

Three different approaches revealed that 41.4% (average value) of the grapevine genome is composed of repetitive/transposable elements (TEs), a slightly higher proportion than that identified in the rice genome, which has a somewhat smaller size². The distribution of repeats and TEs along the chromosomes is quite uneven (see below). All classes and superfamilies of TEs are represented in the grapevine genome, with a large prevalence of class I elements over class II and helitrons (rolling-circle transposons) (Supplementary Table 7). An analysis of the distribution of the repetitive elements in the different fractions of the grapevine genome based on the current annotation shows that introns are quite rich in repeats and TEs (data not shown). In addition, 12.4% of the intron sequence contains transposons as determined using our set of manually annotated elements, most of which (75%) correspond to LINE (long interspersed element) retrotransposons, which therefore seem to have contributed specifically to the intron size observed in grapevine (Supplementary Table 8).

In eukaryotes with large genomes, the coding and repeated elements are distributed over the chromosomes and may be more or less interlaced, hence defining gene-poor and gene-rich regions. It has previously been noticed that the distribution of the genes along the chromosomes of rice and *Arabidopsis thaliana* is fairly homogeneous^{2,3}. In contrast, we observe large regions that alternate between high and low gene density in *V. vinifera* (Supplementary Figs 2 and 3). As expected, the density of TEs reflects a pattern substantially complementary to gene density. We observe a similar characteristic in the genome sequence of poplar, therefore indicating a dynamic for the invasion of TEs that is shared with the grapevine (Supplementary Fig. 3).

A striking feature of the grapevine proteome lies in the existence of large families related to wine characteristics, which have a higher gene copy number than in the other sequenced plants. Stilbene synthases (STSs) drive the synthesis of resveratrol, the grapevine phytoalexin that has been associated with the health benefits associated with moderate consumption of red wine^{15,16}. The family of genes encoding STSs has a noticeable expansion: 43 genes have been identified. Of these, 20 have previously been shown to be expressed after infection by *Plasmopara viticola*, thus confirming that they are likely to be functional. The terpene synthases (TPSs) drive the synthesis of terpenoids; these secondary metabolites are major components of resins, essential oils and aromas (their relative abundance is directly correlated with the aromatic features of wines¹⁷) and are involved in plant–environment interactions. In comparison with the 30–40 genes of this family in *Arabidopsis*, rice and poplar, the grapevine TPS family is more than twice as large, with 89 functional genes and 27 pseudogenes. Classification based on known plant homologues reveals that the subclass of putative monoterpene synthases represents only 15% of the *Arabidopsis* TPS family¹⁸ whereas this subclass represents 40% of the grapevine TPS family. This result suggests a high diversification of grapevine monoterpene synthases that specifically produce C₁₀ terpenoids present in aroma (such as geraniol, linalool, cineole and α -terpineol). Furthermore, the grapevine genome annotation has also revealed genes encoding homologues to the two forms of geranyl diphosphate synthases (GPPSs), the enzymes that produce the substrate for monoterpene synthases: both the

homodimeric GPPS and the heterodimeric form are present; the latter is present only in plants such as *Mentha piperita* and *Clarkia breweri*, which produce large quantities of monoterpenes¹⁹. Most of the STS and TPS genes occur as 20 clusters, including up to 33 paralogous genes located in a 680-kb stretch.

Because global duplication events seem to be a frequent event in plant evolution²⁰, we searched the genome of *V. vinifera* for paralogous regions by using protein sequence similarity. Paralogous regions are defined as chromosome fragments in which homologous genes are present in clusters. Statistical analysis²¹ of these clusters reveals that 94.5% have high probability of being paralogous ($P < 10^{-4}$; Supplementary Table 11). Most *Vitis* gene regions have two different paralogous regions, which we have grouped together as triplets (Supplementary Fig. 5; coverage details in Supplementary Table 10). We conclude that the present-day grapevine haploid genome originated from the contribution of three ancestral genomes. It is yet to be demonstrated whether this content came from a true hexaploidization event or through successive genome duplications. The resulting plant had a diploid content that corresponds to the three full diploid contents of the three ancestors; it may therefore be described as a 'palaeo-hexaploid' organism. A number of rearrangements have affected the original three complements after the formation of the palaeo-hexaploid state. However, the gene order has been sufficiently conserved to permit the alignment of most regions with their two siblings.

We explored the time of formation of the palaeo-hexaploid arrangement by comparing grapevine gene regions with those of other completely sequenced plant genomes. If the palaeo-hexaploid complement is present in another species, it should result in a one-for-one pairing of gene regions between the two species considered. In contrast, if another species's genome evolved before palaeo-hexaploid formation, it should result in a one-to-three relationship between the other species and the grapevine genome. The available genome sequences were those of poplar¹, *Arabidopsis*³ and rice (*Oryza sativa*)², of which poplar is considered to be most closely related to grapevine. All clusters constructed between the orthologues in the three comparisons have $P < 10^{-4}$ (Table 1c). When the gene order in poplar is compared with that in grapevine, there are two clear distributions. First, the grapevine regions align with two poplar segments, as would be expected from a recent whole-genome duplication (WGD) in the poplar lineage¹. Second, each of the three grapevine regions that form a homologous triplet recognizes different pairs of poplar segments (Fig. 1a and Supplementary Fig. 6). This shows that the palaeo-hexaploidy observed in grapevine was already present in its common ancestor with poplar.

Poplar belongs to the Eurosid I clade. The sister clade to Eurosid I is that of Eurosid II, which contains the model species *Arabidopsis*. Its gene order was compared with that in the grapevine genome. Two distributions appear: first, most grapevine regions correspond to four *Arabidopsis* segments (Supplementary Fig. 7); second, each component of a triplicated group in grapevine recognizes four different regions in *Arabidopsis* (Fig. 1b). This shows that the grapevine palaeo-hexaploidy was present in the common ancestor to *Arabidopsis* and grapevine, and therefore that it is a trait common to all Eurosids. This is confirmed by the homology level distribution between paralogues of the grapevine, indicating a lower conservation than between *Vitis/Arabidopsis* orthologues (Supplementary Fig. 4). The Eurosid group contains many economically important flowering plants such as legumes, cotton and Brassicaceae. Our present results establish these species as having a palaeo-hexaploid common ancestor. The grapevine/*Arabidopsis* comparison also reveals that the *Arabidopsis* lineage underwent two WGDs after its separation from the Eurosid I clade^{21–24}. This contradicts some models based on more indirect evidence that placed the most ancient of these two duplications at the base of the Eurosid group, or even earlier^{4,20–22}. Some studies had also suggested a possible third duplication event in the distant past of the *Arabidopsis* lineage, potentially at the base of

the angiosperm radiation. The controversy about this third event is now resolved by the *Vitis* genome comparisons: this event corresponds to the palaeo-hexaploidy formation that remains evident in the grapevine genome but has been difficult to characterize in *Arabidopsis* and poplar because of the more recent WGDs. In particular, the *Arabidopsis* genome lineage has undergone many rearrangements and chromosome fusions such that the ancestral gene order is particularly difficult to deduce from this species (Fig. 2).

Grapevines, like *Arabidopsis* and poplar, are dicotyledonous plants that diverged from monocotyledons about 130–240 Myr ago^{25,26}.

Because rice is a monocotyledon, we assessed the presence or absence of palaeo-hexaploidy in its genome sequence. The observed pattern is the opposite of that seen for *Arabidopsis* and poplar: constituents of a grapevine triplet are generally orthologous to the same group of rice regions (Fig. 1c and Supplementary Fig. 11). Because rice and grapevine are phylogenetically distant, it is more difficult to detect relations of orthology across the two whole genomes: rearrangements, duplication and gene loss have affected the gene orders differently in the two lineages (Supplementary Fig. 10). Even with this limitation, we observed numerous cases of one-to-three relationships between

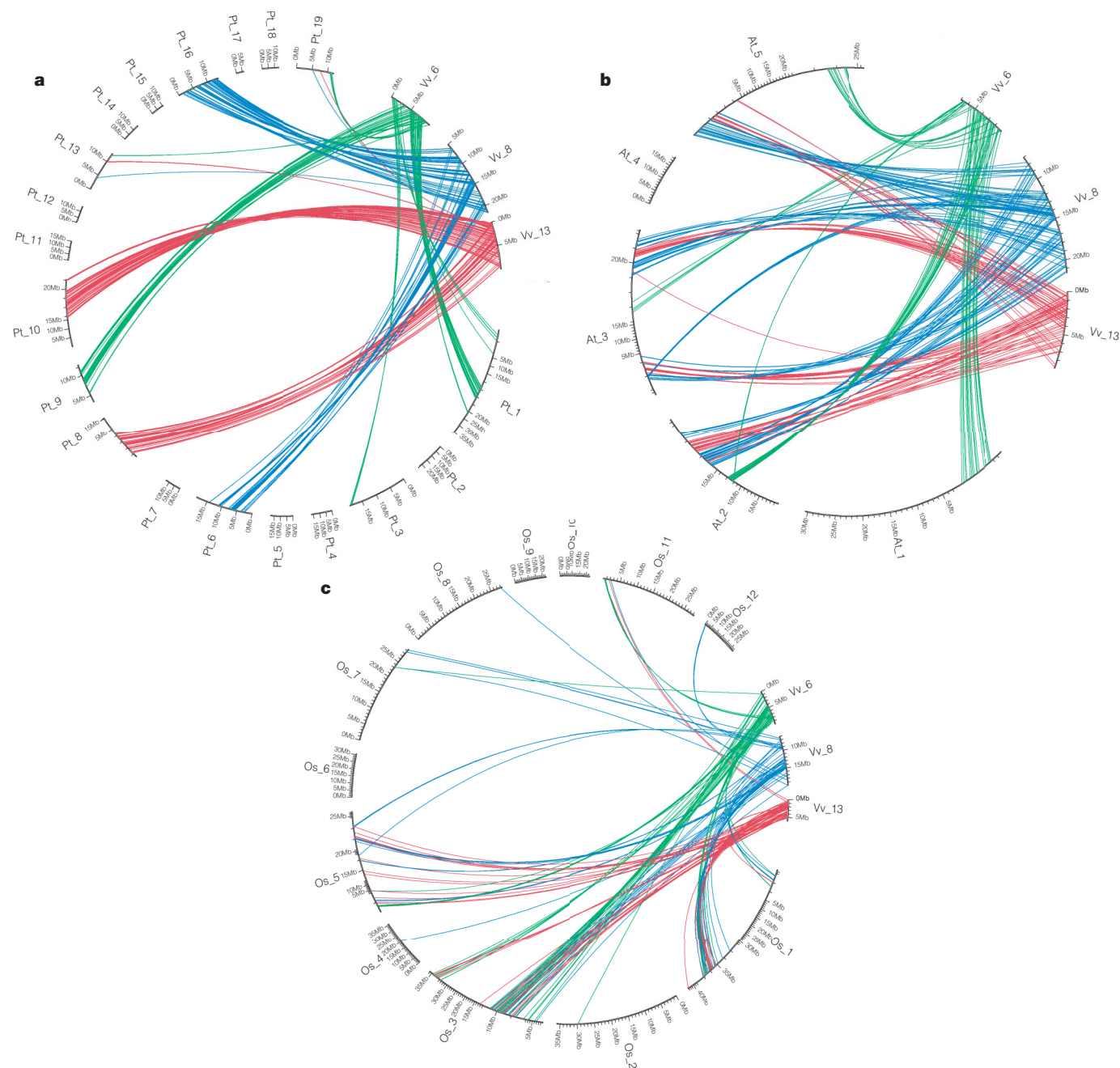


Figure 1 | Comparison between three paralogous *Vitis* genomic regions and their orthologues in *P. trichocarpa*, *A. thaliana* and *O. sativa*. Orthologous gene pairs are joined with a different colour for each of the three paralogous grapevine chromosomes 6 (green), 8 (blue) and 13 (red). **a**, Orthologous regions in the poplar genome are different for each of the three *Vitis* chromosomes, showing that the triplication predates the poplar/*Vitis* separation. One *Vitis* region recognizes two poplar segments because of a WGD in the poplar lineage after the separation. **b**, Orthologous regions with *Arabidopsis* are different for each of the three *Vitis* chromosomes. This

shows that the *Arabidopsis/Vitis* ancestor had the same palaeo-hexaploid content. One *Vitis* region corresponds to four *Arabidopsis* segments, indicating the presence of two WGDs in the *Arabidopsis* lineage after separation from the *Vitis* lineage. **c**, Orthologous regions in rice are the same for the three paralogous chromosomes. This indicates that the triplication was not present in the common ancestor of monocotyledons and dicotyledons. The presence in rice of different homologous blocks is due to global duplications in the rice lineage after divergence from dicotyledons.

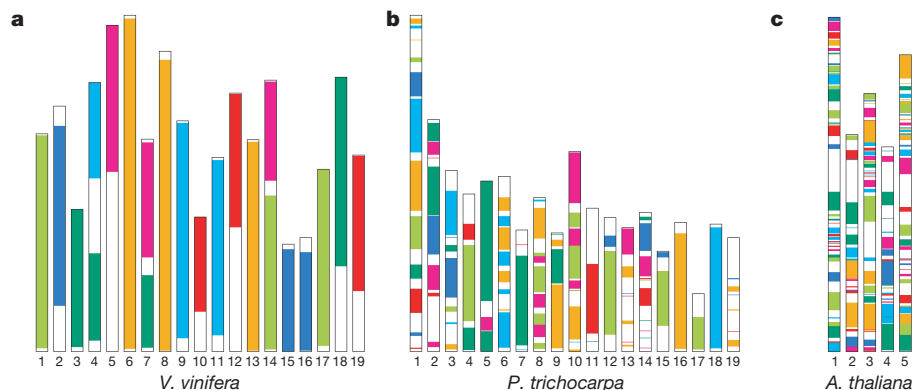


Figure 2 | Schematic representation of paralogous regions derived from the three ancestral genomes in the karyotypes of *V. vinifera*, *P. trichocarpa* and *A. thaliana*. Each colour corresponds to a syntenic region between the three ancestral genomes that were defined by their occurrence as linked clusters in grapevine, independently of intrachromosomal rearrangements.

rice and grapevine (Supplementary Figs 8, 9 and 11); 23% of orthologous blocks include the paralogous regions that originate from the grapevine palaeo-hexaploidy. For *Arabidopsis*, this number is as low as 1.4% (this difference is significant at 5%: $\chi^2 = 8.9$; Supplementary Table 12), despite the fact that the *Arabidopsis* genome has suffered many gene losses since its two WGDs. These gene losses would be expected to obscure the orthologous relations with the grapevine genome, but they are clearly insufficient to explain the high number of one-to-three relationships observed in the rice–grapevine comparison. The most probable explanation for this excess is that the rice ancestor did not exhibit the palaeo-hexaploidy observed in the grapevine, poplar and *Arabidopsis*.

These findings are summarized in Fig. 3: the triplicated arrangement is apparent after the separation of the monocotyledons and dicotyledons and before the spread of the Eurosids clade. Future genome sequencing projects for other clades of dicotyledons, such as Solanaceae or basal eudicots, will help in situating the triplication event more precisely, and eventually in establishing its precise nature (hexaploidization or genome duplications at distant times).

Public access to the grapevine genome sequence will help in the identification of genes underlying the agricultural characteristics of

The *V. vinifera* genome (a) is by far the closest to the ancestral arrangement, whereas that of *Arabidopsis* (c) is thoroughly rearranged, and *P. trichocarpa* (b) presents an intermediate situation. The seven colours probably correspond to linkage groups at the time of the palaeo-hexaploid ancestor.

this species, including domestication traits. A selective amplification of genes belonging to the metabolic pathways of terpenes and tannins has occurred in the grapevine genome, in contrast with other plant genomes. This suggests that it may become possible to trace the diversity of wine flavours down to the genome level. Grapevine is also a crop that is highly susceptible to a large diversity of pathogens including powdery mildew, oidium and Pierce disease. Other *Vitis* species such as *V. riparia* or *V. cinerea*, which are known to be resistant to several of these pathogens, are interfertile with *V. vinifera* and can be used for the introduction of resistance traits by advanced backcrosses²⁷ or by gene transfer. Access to the *Vitis* sequence and the exploitation of synteny will speed up this process of introgression of pathogen resistance traits. As a consequence of this, it is hoped that it will also prompt a strong decrease in pesticide use.

The high quality of the assembly, due mainly to the highly homozygous nature of the PN40024 line, enables the discovery of three ancestral genomes constituting the diploid content of grapevine. The Greek historian Thucydides wrote that Mediterranean people began to emerge from ignorance when they learnt to cultivate olives and grapes. This first characterization of the grapevine genome, with its indication of a palaeo-hexaploid ancestral genome for many dicotyledonous plants, addresses fundamental questions related to the origin and importance of this event in the history of flowering plants. Future work may help in correlating the differential fates of the three gene complements with phenotypic traits of dicotyledonous species.

METHODS SUMMARY

Gene annotation. Protein-coding genes were predicted by combining *ab initio* models, *V. vinifera* complementary DNA alignments, and alignments of proteins and genomic DNA from other species. The integration of the data was performed with GAZE²⁸. Details are given in Supplementary Information.

Paralogous and orthologous gene sets. Statistical testing of homologous regions was performed as described in ref. 21.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 5 April; accepted 7 August 2007.

Published online 26 August 2007.

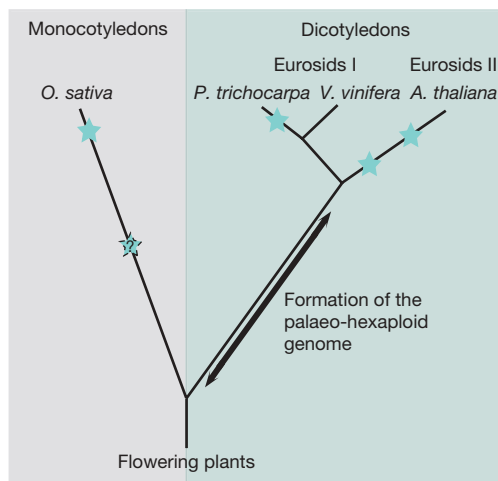


Figure 3 | Positions of the polyploidization events in the evolution of plants with a sequenced genome. Each star indicates a WGD (tetraploidization) event on that branch. The question mark indicates that ancient events are visible in the rice genome that would require other monocotyledon genome sequences to be resolved. The formation of the palaeo-hexaploid ancestral genome occurred after divergence from monocotyledons and before the radiation of the Eurosids.

1. Tuskan, G. A. *et al.* The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596–1604 (2006).
2. International Rice Genome Sequencing Project. The map-based sequence of the rice genome. *Nature* **436**, 793–800 (2005).
3. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815 (2000).
4. De Bodt, S., Maere, S. & Van de Peer, Y. Genome duplication and the origin of angiosperms. *Trends Ecol. Evol.* **20**, 591–597 (2005).
5. Scannell, D. R., Byrne, K. P., Gordon, J. L., Wong, S. & Wolfe, K. H. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature* **440**, 341–345 (2006).

6. Jaillon, O. *et al.* Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431**, 946–957 (2004).
7. Aury, J. M. *et al.* Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature* **444**, 171–178 (2006).
8. Maere, S. *et al.* Modeling gene and genome duplications in eukaryotes. *Proc. Natl Acad. Sci. USA* **102**, 5454–5459 (2005).
9. Blanc, G. & Wolfe, K. H. Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **16**, 1679–1691 (2004).
10. Seoighe, C. & Gehring, C. Genome duplication led to highly selective expansion of the *Arabidopsis thaliana* proteome. *Trends Genet.* **20**, 461–464 (2004).
11. McGovern, P. E., Hartung, U., Badler, V., Glusker, D. L. & Exner, L. J. The beginnings of wine making and viniculture in the ancient Near East and Egypt. *Expedition* **39**, 3–21 (1997).
12. Jaffe, D. B. *et al.* Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome Res.* **13**, 91–96 (2003).
13. Lodhi, M. A., Daly, M. J., Ye, G. N., Weeden, N. F. & Reisch, B. I. A molecular marker based linkage map of *Vitis*. *Genome* **38**, 786–794 (1995).
14. Doligez, A. *et al.* An integrated SSR map of grapevine based on five mapping populations. *Theor. Appl. Genet.* **113**, 369–382 (2006).
15. Baur, J. A. *et al.* Resveratrol improves health and survival of mice on a high-calorie diet. *Nature* **444**, 337–342 (2006).
16. Baur, J. A. & Sinclair, D. A. Therapeutic potential of resveratrol: the *in vivo* evidence. *Nature Rev. Drug Discov.* **5**, 493–506 (2006).
17. Mateo, J. J. & Jimenez, M. Monoterpenes in grape juice and wines. *J. Chromatogr. A* **881**, 557–567 (2000).
18. Aubourg, S., Lecharny, A. & Bohlmann, J. Genomic analysis of the terpenoid synthase (AtTPS) gene family of *Arabidopsis thaliana*. *Mol. Genet. Genomics* **267**, 730–745 (2002).
19. Tholl, D. *et al.* Formation of monoterpenes in *Antirrhinum majus* and *Clarkia breweri* flowers involves heterodimeric geranyl diphosphate synthases. *Plant Cell* **16**, 977–992 (2004).
20. Adams, K. L. & Wendel, J. F. Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* **8**, 135–141 (2005).
21. Simillion, C., Vandepoele, K., Van Montagu, M. C., Zabeau, M. & Van de Peer, Y. The hidden duplication past of *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **99**, 13627–13632 (2002).
22. Bowers, J. E., Chapman, B. A., Rong, J. & Paterson, A. H. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**, 433–438 (2003).
23. Vision, T. J., Brown, D. G. & Tanksley, S. D. The origins of genomic duplications in *Arabidopsis*. *Science* **290**, 2114–2117 (2000).
24. Blanc, G., Hokamp, K. & Wolfe, K. H. A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. *Genome Res.* **13**, 137–144 (2003).
25. Wolfe, K. H., Gouy, M., Yang, Y. W., Sharp, P. M. & Li, W. H. Date of the monocot–dicot divergence estimated from chloroplast DNA sequence data. *Proc. Natl Acad. Sci. USA* **86**, 6201–6205 (1989).
26. Crane, P. R., Friis, E. M. & Pedersen, K. R. The origin and early diversification of angiosperms. *Nature* **374**, 27–33 (1995).
27. Eshed, Y. & Zamir, D. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* **141**, 1147–1162 (1995).
28. Howe, K. L., Chothia, T. & Durbin, R. GAZE: a generic framework for the integration of gene-prediction data by dynamic programming. *Genome Res.* **12**, 1418–1427 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The sequencing of the grapevine genome was launched and carried out after a scientific cooperation agreement between the Ministry of Agriculture in France and the Ministry of Agriculture in Italy, involving l'Institut National de la Recherche Agronomique (INRA), Consiglio per la Ricerca e Sperimentazione in Agricoltura (CRA) and Friuli Venezia Giulia Region. This work

was financially supported by Consortium National de Recherche en Génomique, Agence Nationale de la Recherche, INRA, and by MiPAF (VIGNA-CRA), Friuli Innovazione, Università di Udine, Federazione BCC, Fondazione CRUP, Fondazione Carigo, Fondazione CRT, Vivai Cooperativi Rauscedo, Eurotech, Livio Felluga, Marco Felluga, Venica e Venica, Le Vigne di Zamò (IGA). We thank S. Cure for correcting the manuscript; F. Câmara and R. Guigo for the calibration of the GeneID gene prediction software, and the Centre Informatique National de l'Enseignement Supérieur for computing resources.

Author Information The final assembly and annotation are deposited in the EMBL/Genbank/DBJ databases under accession numbers CU459218–CU462737 (for all scaffolds) and CU462738–CU462772 (for chromosome reconstitutions and unanchored scaffolds). An annotation browser and further information on the project are available from <http://www.genoscope.cns.fr/vitis>, <http://www.vitisgenome.it/> and <http://www.appliedgenomics.org/>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.W. (pwinner@genoscope.cns.fr).

The French-Italian Public Consortium for Grapevine Genome Characterization

Olivier Jaillon^{1*}, Jean-Marc Aury^{1*}, Benjamin Noel¹, Alberto Policriti^{2,3}, Christian Clepet⁴, Alberto Casagrande^{2,5}, Nathalie Choise^{1,4}, Sébastien Aubourg⁴, Nicola Vitulo^{6,15}, Claire Jubin¹, Alessandro Vezzi^{6,15}, Fabrice Legeai⁷, Philippe Huguene⁸, Corinne Dasilva¹, David Horner^{9,15}, Erica Mica^{9,15}, Delphine Jublot⁴, Julie Poulain¹, Clémence Bruyère⁴, Alain Billault¹, Béatrice Segurens¹, Michel Gouyvenoux¹, Edgardo Ugarte¹, Federica Cattonaro², Véronique Anthouard¹, Virginie Vico¹, Cristian Del Fabbro^{2,3}, Michaël Alaux⁷, Gabriele Di Gasparo^{2,5}, Vincent Dumas⁸, Nicoletta Felice^{2,5}, Sophie Paillard⁴, Irena Juman^{2,5}, Marco Moroldo⁴, Simone Scalabrini^{2,3}, Aurélie Canaguier⁴, Isabelle Le Clainche⁴, Giorgio Malacrida^{6,15}, Eléonore Durand⁷, Graziano Pesole^{10,11,15}, Valérie Laucou¹², Philippe Chatelet¹³, Didier Merdinoglu⁸, Massimo Delledonne^{14,15}, Mario Pezzotti^{15,16}, Alain Lecharny⁴, Claude Scarpelli¹, François Artiguenave¹, M. Enrico Pè^{9,15}, Giorgio Valle^{6,15}, Michele Morgante^{2,5}, Michel Caboche⁴, Anne-Françoise Adam-Blondon⁴, Jean Weissenbach¹, Francis Quétier¹ & Patrick Wincker¹

*These authors contributed equally to this work.

Affiliations for participants: ¹Genoscope (CEA) and UMR 8030 CNRS-Genoscope-Université d'Evry, 2 rue Gaston Crémieux, BP5706, 91057 Evry, France. ²Istituto di Genomica Applicata, Parco Scientifico e Tecnologico di Udine, Via Linussio 51, 33100 Udine, Italy. ³Dipartimento di Matematica ed Informatica, Università degli Studi di Udine, via delle Scienze 208, 33100 Udine, Italy. ⁴URGV, UMR INRA 1165, CNRS-Université d'Evry Génomique Végétale, 2 rue Gaston Crémieux, BP5708, 91057 Evry cedex, France. ⁵Dipartimento di Scienze Agrarie ed Ambientali, Università degli Studi di Udine, via delle Scienze 208, 33100 Udine, Italy. ⁶CRIBI, Università degli Studi di Padova, viale G. Colombo 3, 35121 Padova, Italy. ⁷URGI, UR1164 Génomique Info, 523, Place des Terrasses, 91034 Evry Cedex, France. ⁸UMR INRA 1131, Université de Strasbourg, Santé de la Vigne et Qualité du Vin, 28 rue de Herrlisheim, BP20507, 68021 Colmar, France. ⁹Dipartimento di Scienze Biomolecolari e Biotecnologie, Università degli Studi di Milano, via Celoria 26, 20133 Milano, Italy. ¹⁰Dipartimento di Biochimica e Biologia Molecolare, Università degli Studi di Bari, via Orabona 4, 70125 Bari, Italy. ¹¹Istituto Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, via Amendola 122/D, 70125 Bari, Italy. ¹²UMR INRA 1097, IRD-Montpellier SupAgro-Univ. Montpellier II, Diversité et Adaptation des Plantes Cultivées, 2 Place Pierre Viala, 34060 Montpellier Cedex 1, France. ¹³UMR INRA 1098, IRD-Montpellier SupAgro-CIRAD, Développement et Amélioration des Plantes, 2 Place Pierre Viala, 34060 Montpellier Cedex 1, France. ¹⁴Dipartimento Scientifico e Tecnologico, Università degli Studi di Verona Strada Le Grazie 15 – Ca' Vignal, 37134 Verona, Italy. ¹⁵Dipartimento di Scienze, Tecnologie e Mercati della Vite e del Vino, Università degli Studi di Verona, via della Pieve, 70 37029 S. Floriano (VR), Italy. ¹⁶VIGNA-CRA Initiative; Consorzio Interuniversitario Nazionale per la Biologia Molecolare delle Pianta, c/o Università degli Studi di Siena, via Banchi di Sotto 55, 53100 Siena, Italy.

METHODS

Genome sequencing. The *V. vinifera* PN40024 genome was sequenced with the use of a whole-genome shotgun strategy. All data were generated by paired-end sequencing of cloned inserts using Sanger technology on ABI3730xl sequencers. Supplementary Table 2 gives the number of reads obtained per library.

Genome assembly and chromosome anchoring. All reads were assembled with Arachne¹². We obtained 20,784 contigs that were linked into 3,830 supercontigs of more than 2 kb. The contig N₅₀ was 64 kb, and the supercontig N₅₀ was 1.9 Mb. The total supercontig size was 498 Mb, remarkably close to the expected size of 475 Mb. This indicates that the PN40024 has retained few heterozygous regions. Remaining heterozygosity was assessed by aligning all supercontigs with each other. We first selected the supercontigs more than 30 kb in size that were covered over more than 40% of their length by another supercontig with more than 95% identity. After visual inspection of the alignments, we added to this list the supercontigs more than 10 kb in size that aligned at more than 40% of their length with supercontigs identified previously. All potential cases were then inspected visually to discard potential heterozygous regions (aligning relatively homogeneously across their complete length) and retained repeated regions (with more heterogeneous alignments). This treatment identified 11 Mb of potentially allelic supercontigs. We confirmed that in most cases their coverage was about half the average of the homozygous supercontigs. Only one supercontig of each allelic pair was therefore conserved in the final assembly, which consists of 3,514 supercontigs (N₅₀ = 2 Mb) containing 19,577 contigs (N₅₀ = 66 kb), totalling 487 Mb. If the haploid genome size of 475 Mb is considered correct, then our final assembly contains only about 12 Mb of remaining heterozygosity, or 2.6%.

A set of 30,151 bacterial artificial chromosome (BAC) fingerprints of the BAC clones of a Cabernet-Sauvignon library²⁹ were assembled into 1,763 contigs with FPC³⁰, v. 8. In parallel, 1,981 markers were anchored on a subset of BAC clones³¹, among which 388 markers mapped onto the genetic map, and 77,237 BAC end sequences were obtained³¹. Blat³² alignments (90% identity on 80% of the length, fewer than five hits) were performed with BAC end sequences on the 3,830 supercontigs of sequences with lengths over 2 kb. The results were then filtered with homemade Perl scripts to keep only the occurrences in which two paired ends were matching at a distance of less than 300 kb and with a consistent orientation. Two supercontigs were considered linked to each other if two BAC links could be found or one BAC link and a BAC contig link. A total number of 111 ultracontigs were constructed with this procedure.

Genome annotation. Several resources were used to build *V. vinifera* gene models automatically with GAZE²⁸. We used predictions of repetitive regions by repeatscout³³, conserved coding regions predicted by the exofish method^{34,35}, genewise³⁶ alignments of proteins from Uniprot³⁷, Geneid³⁸ and Snap³⁹ *ab initio* gene predictions, and alignments of several cDNA resources (Supplementary Information).

A weight was assigned to each resource to further reflect its reliability and accuracy in predicting gene models. This weight acts as a multiplier for the score of each information source, before being processed by GAZE. When applied to the entire assembled sequence, GAZE predicted 30,434 gene models.

Paralogous and orthologous gene sets. We identified orthologous genes in six pairs of genomes from four species: *A. thaliana*, *O. sativa*, *P. trichocarpa* and *V. vinifera*. Each pair of predicted gene sets was aligned with the Smith–Waterman algorithm, and alignments with a score higher than 300 (BLOSUM62; gapo = 10, gape = 1) were retained. Two genes, A from genome GA and B from genome GB, were considered orthologues if B was the best match for gene A in GB and A was the best match for B in GA.

For each orthologous gene set with *V. vinifera*, clusters of orthologous genes were generated. A single linkage clustering with a euclidean distance was used to group genes. The distances were calculated with the gene index in each chromosome rather than the genomic position. The minimal distance between two orthologous genes was adapted in accordance with the selected genomes. Finally, we retained only clusters that were composed of at least six genes for *Arabidopsis* and *O. sativa*, and eight genes for *P. trichocarpa* (Supplementary Table 10).

To validate the clustering quality we used a method described previously²¹. For each cluster we computed the probability of finding this cluster in the gene homology matrix (Supplementary Table 11). This matrix was constructed from two compared chromosomes with genes numbered according to their position on each chromosome, with no reference to physical distances.

Paralogous genes were computed by comparing all-against-all of *V. vinifera* proteins by using blastp, and alignments with an expected value of less than 0.1 were retained and realigned with the Smith–Waterman algorithm⁴⁰. Two genes A and B were considered paralogues if B was the best match for gene A and A was the best match for B. Moreover, clusters of paralogous genes were constructed in the same fashion as orthologous clusters (Supplementary Table 10).

29. Adam-Blondon, A. F. *et al.* Construction and characterization of BAC libraries from major grapevine cultivars. *Theor. Appl. Genet.* **110**, 1363–1371 (2005).
30. Soderlund, C., Humphray, S., Dunham, A. & French, L. Contigs built with fingerprints, markers, and FPC V4.7. *Genome Res.* **10**, 1772–1787 (2000).
31. Lamoureaux, D. *et al.* Anchoring of a large set of markers onto a BAC library for the development of a draft physical map of the grapevine genome. *Theor. Appl. Genet.* **113**, 344–356 (2006).
32. Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
33. Price, A. L., Jones, N. C. & Pevzner, P. A. *De novo* identification of repeat families in large genomes. *Bioinformatics* **21** (Suppl. 1), i351–i358 (2005).
34. Roest Crolius, H. *et al.* Estimate of human gene number provided by genome-wide analysis using *Tetradon nigroviridis* DNA sequence. *Nature Genet.* **25**, 235–238 (2000).
35. Jaillon, O. *et al.* Genome-wide analyses based on comparative genomics. *Cold Spring Harb. Symp. Quant. Biol.* **68**, 275–282 (2003).
36. Birney, E., Clamp, M. & Durbin, R. GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004).
37. Bairoch, A. *et al.* The Universal Protein Resource (UniProt). *Nucleic Acids Res.* **33**, D154–D159 (2005).
38. Parra, G., Blanco, E. & Guigo, R. GeneID in *Drosophila*. *Genome Res.* **10**, 511–515 (2000).
39. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
40. Smith, T. F. & Waterman, M. S. Identification of common molecular subsequences. *J. Mol. Biol.* **147**, 195–197 (1981).

LETTERS

Genetic variation in a human odorant receptor alters odour perception

Andreas Keller^{1*}, Hanyi Zhuang^{2*}, Qiuyi Chi², Leslie B. Vosshall¹ & Hiroaki Matsunami^{2,3}

Human olfactory perception differs enormously between individuals, with large reported perceptual variations in the intensity and pleasantness of a given odour. For instance, androstenone (5 α -andro-16-en-3-one), an odorous steroid derived from testosterone, is variously perceived by different individuals as offensive (“sweaty, urinous”), pleasant (“sweet, floral”) or odourless^{1–3}. Similar variation in odour perception has been observed for several other odours^{4–6}. The mechanistic basis of variation in odour perception between individuals is unknown. We investigated whether genetic variation in human odorant receptor genes accounts in part for variation in odour perception between individuals^{7,8}. Here we show that a human odorant receptor, OR7D4, is selectively activated *in vitro* by androstenone and the related odorous steroid androstadienone (androsta-4,16-dien-3-one) and does not respond to a panel of 64 other odours and two solvents. A common variant of this receptor (OR7D4 WM) contains two non-synonymous single nucleotide polymorphisms (SNPs), resulting in two amino acid substitutions (R88W, T133M; hence ‘RT’) that severely impair function *in vitro*. Human subjects with RT/WM or WM/WM genotypes as a group were less sensitive to androstenone and androstadienone and found both odours less unpleasant than the RT/RT group. Genotypic variation in OR7D4 accounts for a significant proportion of the valence (pleasantness or unpleasantness) and intensity variance in perception of these steroidal odours. Our results demonstrate the first link between the function of a human odorant receptor *in vitro* and odour perception.

We investigated the hypothesis that polymorphisms in odorant receptors contribute to variability in human odour perception by combining a cell-based assay technique to identify active ligands for odorant receptors⁹ with an olfactory psychophysical study of a diverse population of human subjects¹⁰. A total of 66 odourants were used to measure both odorant receptor responses *in vitro* and psychophysical responses in human subjects, with a focus on odorous steroids because the perception of these odours is exceptionally variable^{3,11}. We cloned a panel of 335 putative human odorant receptors, representing more than 85% of the odorant receptors with full open reading frames, and expressed them in Hana3A cells, an HEK293T-derived cell line stably expressing accessory factors for odorant receptor expression^{9,12}. We screened for androstenone-mediated stimulation with a luciferase reporter⁹. Among the receptors tested, that encoded by OR7D4 showed the strongest responses to androstenone (Fig. 1a). Recent expression analysis shows that OR7D4 is selectively expressed in human nasal epithelium¹³. Several other receptors showed smaller responses to androstenone that may be relevant *in vivo* (Supplementary Table 1), but these were not investigated further. The failure of a specific odorant receptor to respond in this assay must be interpreted with caution because it may reflect a failure of the odorant receptor to be functional in the assay rather than a lack of sensitivity to the tested odour.

A search for polymorphisms in OR7D4 in SNP databases and sequencing the coding region of OR7D4 in 391 subjects identified 13 non-synonymous SNPs in this receptor, with four occurring at

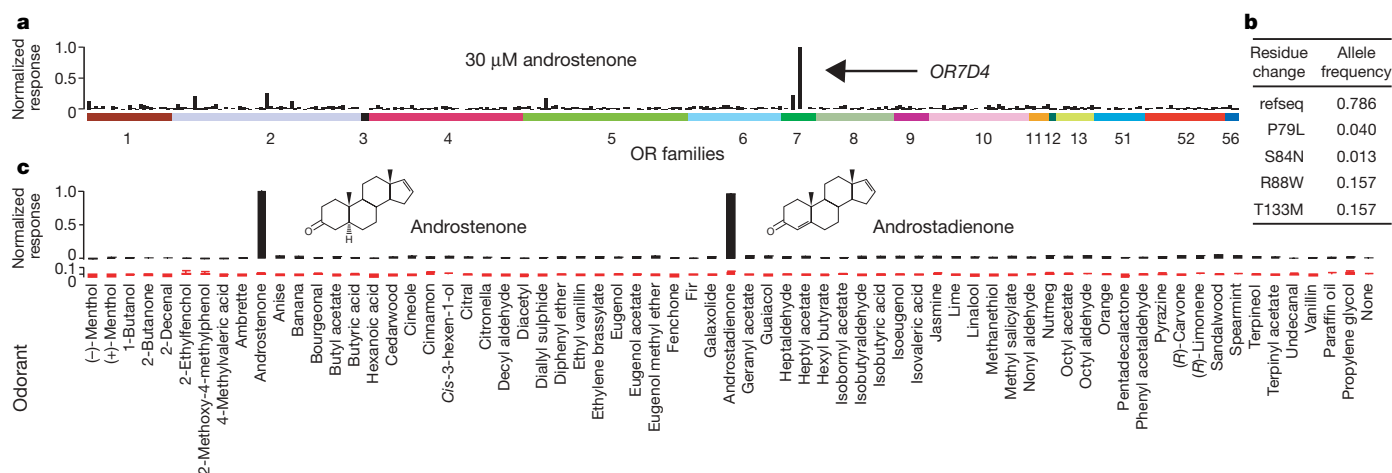


Figure 1 | OR7D4 is selectively activated by androstenone and androstadienone. **a**, Hana3A cell luciferase assays of 335 unique human odorant receptors; the concentration of androstenone used was 30 μ M. Numbers and coloured bars indicate different OR families. **b**, Allele

frequencies of common variants. refseq, reference sequence. **c**, OR7D4 RT (black columns) and WM (red columns) tested against 66 odours and 2 solvents (30 μ M, or 1/30,000). Normalized responses are shown as means and s.e.m. ($n = 4$).

¹Laboratory of Neurogenetics and Behaviour, The Rockefeller University, 1230 York Avenue, New York, New York 10065, USA. ²Department of Molecular Genetics and Microbiology, and ³Department of Neurobiology, Duke University Medical Centre, Research Drive, Durham, North Carolina 27710, USA.

*These authors contributed equally to this work.

frequencies greater than 1% (Supplementary Table 2). Two non-synonymous polymorphisms in complete linkage disequilibrium in this population occurred at the highest frequency and led to two amino-acid changes (R88W and T133M; Fig. 1b). We refer to the most common allele of this receptor, or the reference sequence, as RT, and to the other as WM.

We investigated the ligand specificity of RT and WM receptor variants *in vitro* with a panel of 66 odours and 2 solvents. OR7D4 RT responds selectively to androstene and the closely related odorous steroid androstadiene but shows no response to any other stimuli tested (Fig. 1c, top). OR7D4 WM shows no response to any compound at the concentrations tested here (Fig. 1c, bottom). Dose-response curves with RT and WM show that the paired SNPs in the WM variant, which affect amino acids in extracellular loop 2 and intracellular loop 2 (Fig. 2a), severely impair function (Fig. 2b). We generated odorant receptors with each of the SNPs and found that OR7D4 R88W and OR7D4 T133M retained an intermediate level of function, suggesting that both residues are important for OR7D4 function (Fig. 2b).

We examined two other SNP variants found at frequencies greater than 1%, which led to amino acid changes P79L and S84N,

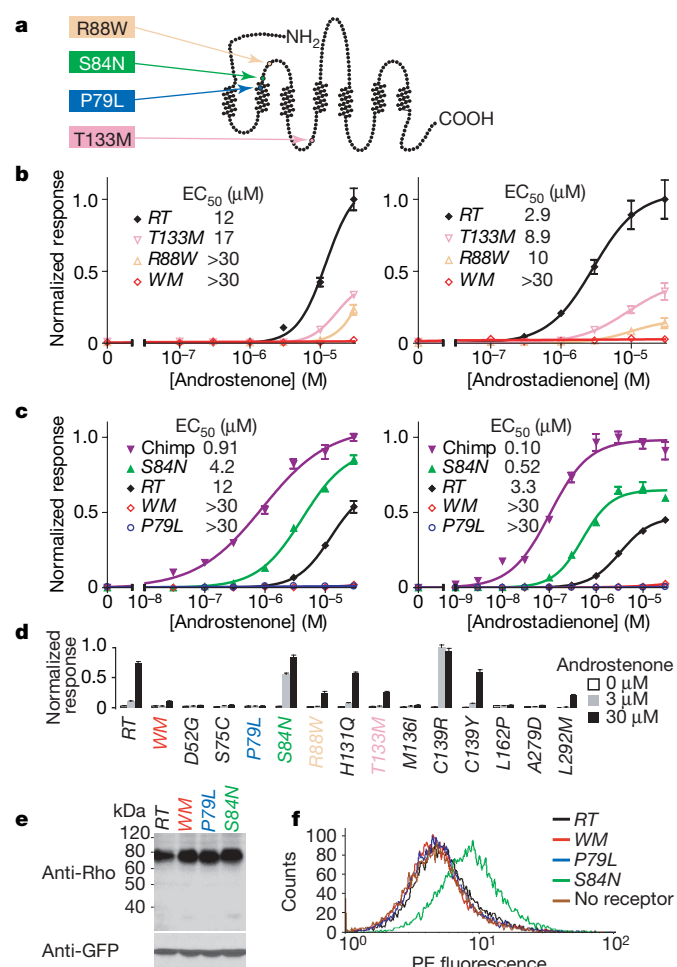


Figure 2 | Functional characterization of OR7D4 polymorphisms. **a**, OR7D4 snake plot with amino acid changes indicated. **b**, **c**, Dose-response curves and EC_{50} values of OR7D4 RT, WM, R88W and T133M (**b**) and of OR7D4 RT, WM, P79L and S84N and chimpanzee OR7D4 (**c**) to androstene (left) and androstadiene (right). **d**, Activity of 13 SNP variants compared with that of RT and WM variants. In **b–d**, normalized responses are shown as means and s.e.m. ($n = 4–6$). **e**, Western blot analysis of whole-cell lysates from HEK293T cells transfected with OR7D4 RT, WM, P79L or S84N and co-transfected with green fluorescent protein (GFP). **f**, Flow cytometry analysis of cell-surface OR7D4 RT, WM, P79L and S84N expression as measured by intensity of phycoerythrin (PE) signal among GFP-positive cells.

respectively (Fig. 1b and 2a). P79L and S84N possess residues R88 and T133 and are referred to by the single variant residue. Analysis of P79L function *in vitro* showed severely impaired function at all concentrations of either steroidal odour tested (Fig. 2c). In contrast, OR7D4 S84N showed remarkable sensitivity to both odours *in vitro*, exceeding the activity of the common RT variant at every concentration tested, with a concentration giving half-maximal response (EC_{50}) to androstadiene less than one-fifth of that of the RT variant (Fig. 2c). The functional differences between these two variants are not due to the amino-terminal epitope tag (Supplementary Fig. 1). The other non-synonymous substitutions in OR7D4 result in varying receptor functions (Fig. 2d).

OR7D4 is situated in a cluster of seven intact odorant receptor genes, but we found that none of the polymorphisms of the six intact odorant receptors in the OR7D4 gene cluster showed significant linkage with the OR7D4 SNPs (Supplementary Fig. 2a and Supplementary Table 3). We tested the responses of all the major variants of odorant receptors in the OR7D4 cluster and found that none showed responses to androstene and androstadiene exceeding that of the impaired WM variant (Supplementary Fig. 2b, c).

The chimpanzee OR7D4 orthologue differs from the human RT reference sequence at five amino acid residues: at the S84N substitution also found in humans and at four additional non-synonymous substitutions not found in humans (V26I, G171V, G227R and K232E). A dose-response analysis of the chimpanzee OR7D4 orthologue *in vitro* showed robust responses to both steroidal odours, exceeding the activity of the human S84N variant (Fig. 2c; compare purple and green curves).

What accounts for the functional differences between OR7D4 variants? We found no obvious difference in subcellular distribution or expression level in permeabilized Hana3A cells expressing RT, WM, P79L or S84N (Supplementary Fig. 3). Western blot analysis confirms that all are expressed at comparable levels (Fig. 2e), and RT, WM and P79L have similar low levels of surface staining as measured by flow cytometry of live cells stained to reveal the N-terminal epitope (Fig. 2f). S84N showed considerably more surface expression (Fig. 2f), suggesting that the increased function of this variant may stem from enhanced stability at the cell surface or from enhanced cell-surface trafficking.

We next asked whether variation in OR7D4 is correlated with variation in the perception of androstene and androstadiene measured in human subjects. Psychophysical data on 391 subjects performing three different tasks were collected: subjects rated the perceived intensity and valence of 66 different odours at two concentrations (Supplementary Fig. 4); detection thresholds were measured to androstene and androstadiene in a subset of subjects, and to three control odours in all subjects^{14,15} (Supplementary Fig. 5); subjects profiled four odours with 146 semantic labels^{10,16} (see Supplementary Methods).

Psychophysical data on the subjects were subsequently divided according to genotype and assessed for the influence of OR7D4 genotype on perceptual phenotype (Supplementary Table 2). Of the 66 odours and two solvents rated by RT/RT and RT/WM subjects, only androstene and androstadiene showed a significant effect of genotype (Fig. 3a and Supplementary Fig. 6). The steroids were rated as less intense by the RT/WM group (Fig. 3a); the proportion of RT/WM subjects rating the high concentration of androstene as “extremely weak” was fourfold that of RT/RT subjects (Supplementary Fig. 6). This phenotype was specific for these two compounds, as the perception of all the other odours was not affected by OR7D4 genotype (Fig. 3a, b). The significant effect of OR7D4 genotype on steroidal odour intensity perception was replicated in both males and females (Supplementary Fig. 7) and in the largest racial category, Caucasians (Supplementary Fig. 8). Although the WM allele strongly affected androstene intensity perception even in heterozygous subjects, the group of WM/WM subjects showed an even stronger effect on intensity ratings, rating both steroidal odours as less intense

than the group of *RT/RT* subjects (Fig. 3b). The few subjects among our population carrying the less frequent *P79L* and *S84N* variants were examined, and *RT/P79L* subjects as a group showed a trend to perceive both androstenone and androstadienone as less intense than the group of *RT/RT* subjects (Fig. 3c). Conversely, *RT/S84N* subjects as a group showed a trend to rate both odours as more intense (Fig. 3c). However, these differences measured in subjects with the rare alleles were not statistically significant.

Detection thresholds of 121 subjects were determined for both steroidal odours (Fig. 3d, e). *RT/WM* subjects as a group had a higher detection threshold—and were therefore less sensitive—to both compounds than the group of *RT/RT* subjects (Fig. 3d) but had normal thresholds to three control odours (data not shown). The threshold for androstenone in the *RT/WM* group was 11-fold, and that for androstadienone 16-fold, of that in the *RT/RT* group. In addition, 46% of *RT/WM* subjects but only 28% of *RT/RT* subjects could not detect the highest concentration of androstenone we provided ($P < 0.05$; χ^2 test). Detection thresholds were also obtained from the few *RT/P79L*, *WM/WM*, *WM/P79L* and *RT/S84N* subjects (Fig. 3e and Supplementary Fig. 9). The proportion of *RT/P79L* subjects unable to detect the highest concentration of androstenone and androstadienone provided was more than double that of *RT/RT* subjects ($P < 0.05$; χ^2 test) (Fig. 3e). It is unclear why the loss of one functional allele has such a profound effect on the median detection threshold. The same trend was found with the *WM/WM* and *WM/P79L* groups (Supplementary Fig. 9), whereas the *RT/S84N* group was more sensitive to both steroids, with lower detection thresholds

than the group of *RT/RT* controls. However, these differences were not statistically significant (Fig. 3e).

We next examined whether variation in *OR7D4* affects the perception of androstenone and androstadienone odour quality. The *RT/WM* group rated both steroidal odours as less unpleasant than the *RT/RT* group (Fig. 4a, b and Supplementary Fig. 10), such that the proportion of *RT/WM* subjects rating the high concentration of androstadienone as “extremely unpleasant” was less than half of that of *RT/RT* subjects (Supplementary Fig. 10). None of the other 64 odours or the solvents showed a statistically significant difference between the genotypes (Fig. 4a), and the effect was statistically significant for both steroidal odours in both males and females (Supplementary Fig. 7). The group of subjects carrying the impaired *RT/P79L* variant showed a trend to rate both androstenone and androstadienone as less unpleasant than *RT/RT* controls (Fig. 4c), whereas the opposite was found in the *RT/S84N* group carrying a more sensitive variant of *OR7D4* (Fig. 4c); however these differences were not statistically significant.

We asked subjects to assess androstenone odour quality by profiling this odour with a standard set of 146 semantic descriptors (see Supplementary Methods)^{10,16}. All descriptors used by more than 10% of the subjects were analysed and descriptor usage in individuals with different genotypes was compared. Of the 74 descriptors used for androstenone, pentadecalactone, vanillin and the solvent propylene glycol, only five differed significantly by genotype (see Supplementary Methods). *OR7D4 RT/WM* subjects were more likely to rate vanillin as smelling “honey”, “sweet” and “vanilla” (Fig. 4d, e),

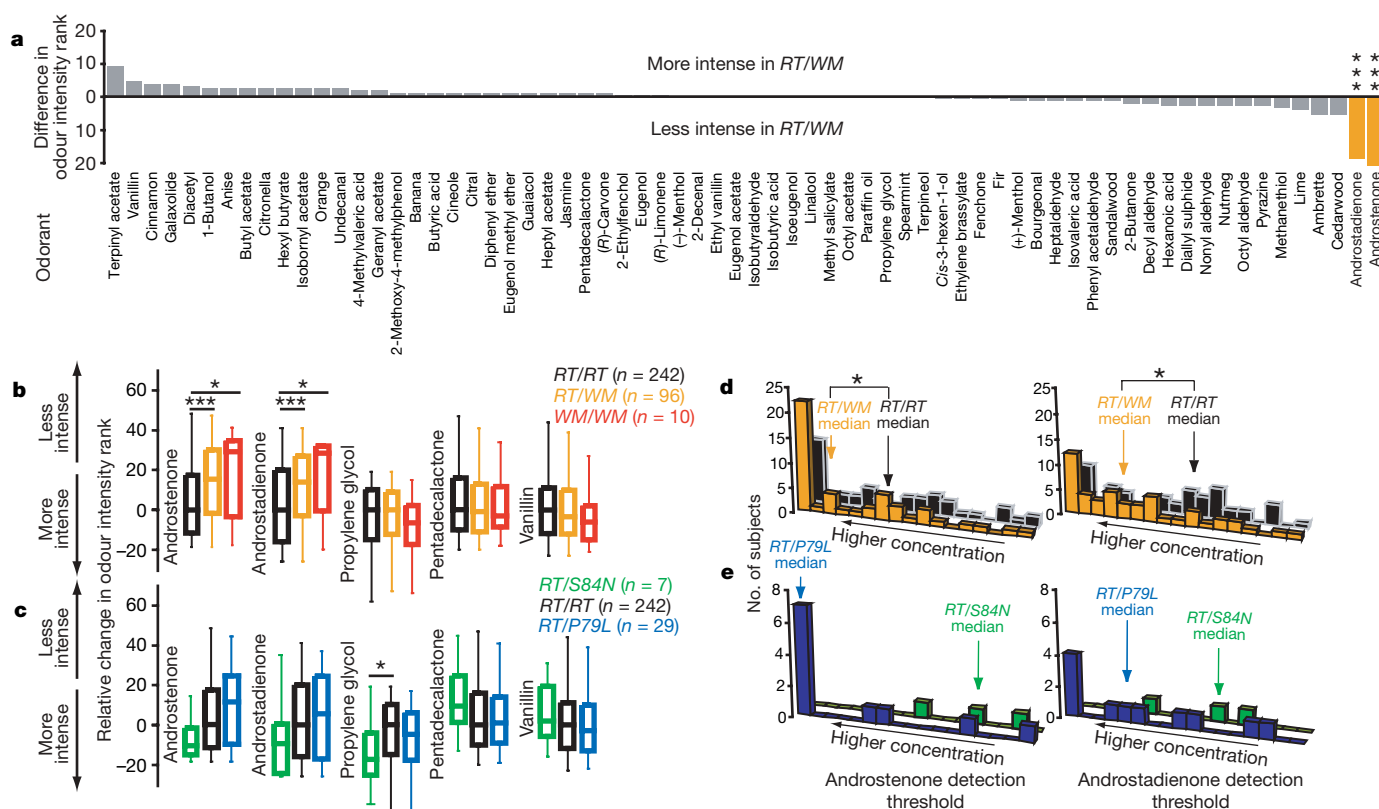


Figure 3 | *OR7D4* variation affects androstenone and androstadienone intensity perception. **a**, Differences in median odour intensity ranking of 66 odours and 2 solvents between *OR7D4 RT/WM* and *RT/RT* groups. Data for two different odour concentrations were pooled. **b**, Change in odour intensity ranking relative to solvent of four odours for the *RT/RT*, *RT/WM* and *WM/WM* groups (**b**) and for the *RT/RT*, *RT/P79L* and *RT/S84N* groups (**c**). The whisker plots show the median rank (normalized to the median rank of the *RT/RT* group), the first and third quartile and the upper and lower limits. Significance was assessed in **a–c** with a Mann–Whitney *U*-test with a Bonferroni correction. Before correction: asterisk, $P < 0.00073$; two

asterisks, $P < 0.00014$; three asterisks, $P < 1.47 \times 10^{-5}$. After correction: asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks, $P < 0.001$.

d, e, Detection thresholds measured in *RT/RT* (n = 47) and *RT/WM* subjects (n = 49) (**d**) and in *RT/P79L* (n = 12) and *RT/S84N* (n = 3) subjects (**e**) plotted as the number of subjects detecting the odour at a given binary dilution (x-axis concentrations are binned from left to right: the first bar represents binary dilution 6, the subsequent 15 bars represent binary dilutions 7–21, and the last bar represents dilutions 22–27). Significance was assessed with a Mann–Whitney *U*-test (asterisk, $P < 0.05$).

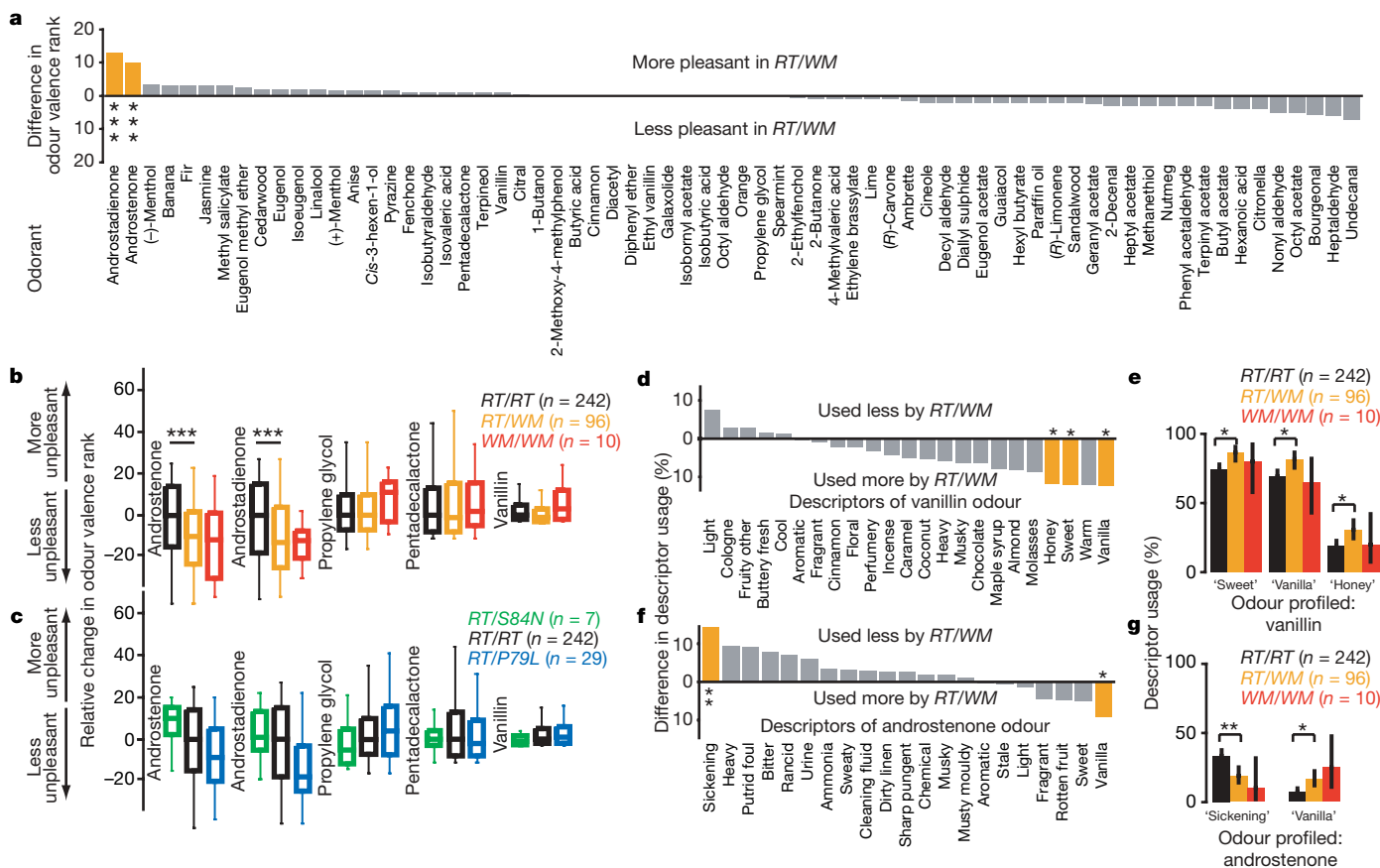


Figure 4 | *OR7D4* variation affects androstenone and androstadienone quality perception. **a**, Differences in median odour valence ranking for the same odours and genotypes as in Fig. 3a. **b**, **c**, Change in odour valence ranking for the same odours and genotypes as in Fig. 3b, c. Significance in **a–c** was assessed with a Mann–Whitney *U*-test with Bonferroni correction. Before correction: asterisk, $P < 0.00073$; two asterisks, $P < 0.00014$; three asterisks, $P < 1.47 \times 10^{-5}$. After correction: asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks, $P < 0.001$. The whisker plots show the median rank (normalized to the median rank of the *RT/RT* group), the first and third quartile and the upper and lower limits. **d**, **e**, Odour profiling of vanillin by *RT/RT* ($n = 242$), *RT/WM* ($n = 96$) and *WM/WM* ($n = 10$) subjects. Plotted

are the differences in descriptor usage by genotype of the 23 descriptors used for vanillin in more than 10% of sessions (**d**) and the percentage of sessions (with 95% confidence intervals) in which the three descriptors that showed significant differences were used (**e**). **f**, **g**, Odour profiling of androstenone for the same genotypes as in **d** and **e**; 21 descriptors were used in more than 10% of all sessions, and two descriptors showed significant differences. Significance was assessed in **d–g** with a χ^2 test with Bonferroni correction. Before correction: asterisk, $P < 0.0022$ (vanillin) and $P < 0.0024$ (androstenone); two asterisks, $P < 0.0004$ (vanillin) and $P < 0.0005$ (androstenone). After correction: asterisk, $P < 0.05$; two asterisks, $P < 0.01$.

42% less likely to consider androstenone “sickening” and 129% more likely to rate it as smelling like “vanilla” than *RT/RT* subjects (Fig. 4f, g). There was no effect of *OR7D4* genotype on odour profiling of pentadecalactone or propylene glycol (data not shown).

Last, we performed a non-parametric regression analysis¹⁷ to estimate the fraction of the phenotypic variance in steroid odour perception attributable to *OR7D4*. We found that *OR7D4* genotype in our population explained 19% and 39% of the variance in the valence and intensity ratings of the steroid odours, respectively ($P < 0.0001$; see Supplementary Methods).

We identify *OR7D4* as a significant heritable factor influencing androstenone and androstadienone perception, thus providing the first reported link between genetic polymorphisms in an odorant receptor gene and altered perception of the ligands that activate this receptor. As predicted by the theory of combinatorial coding¹⁸, we find that polymorphisms in the *OR7D4* protein-coding sequence alone do not fully account for specific anosmia to androstenone and androstadienone. We think it likely that additional human odorant receptors sensitive to androstenone and androstadienone remain to be discovered.

Previous work indicated that sensitivity to androstenone^{19,20}, and to several other odours²¹, is modulated by non-genetic effects such as central processing²² or peripheral sensitization²³ that might obscure any underlying genetic influences on androstenone perception.

About half of the subjects initially unable to detect androstenone were able to smell this compound after being exposed to androstenone daily for six weeks¹⁹. Although it has been assumed that olfactory induction is a phenomenon that occurs only in individuals with specific anosmia to androstenone, more recent studies have found that induction of enhanced olfactory sensitivity seems to be a general phenomenon affecting several odours²¹. The role of *OR7D4* in this sensitization can now be tested in subjects chronically exposed to androstenone.

In this study we investigated only the olfactory percept reported when odorous steroids were sniffed, but olfactory exposure to androstenone and androstadienone has also been shown to induce several physiological responses in both men and women^{24,25}. The identification of an odorant receptor gene that is strongly correlated with the perception of these odours will permit future analysis of olfactory-induced autonomic responses in humans.

METHODS SUMMARY

Heterologous expression of human odorant receptors. A total of 423 human odorant receptors, including 335 predicted functional receptors, were cloned. The chimpanzee *OR7D4* orthologue was cloned from chimpanzee genomic DNA (Coriell Cell Repositories). Odorant receptors containing the first 20 amino acids of human rhodopsin²⁶ in pCI (Promega) were expressed in the Hana3A cell line together with a short form of *mRTP1* called *RTP1S* (M37 to

the carboxy-terminal end), which enhances functional expression of the odorant receptors¹². For immunocytochemistry, cells were fixed, permeabilized and incubated with monoclonal anti-rhodopsin antibody (4D2; ref. 27), followed by Cy3-conjugated donkey anti-mouse IgG (Jackson Immunologicals). For fluorescence-activated cell sorting analysis, this antibody was conjugated with phycoerythrin.

Human odorant receptor genotyping and sequencing. Venous blood was collected from all subjects, and genomic DNA was prepared with the Qiagen PAXgene blood DNA kit. Polymorphisms in *OR7D4* were assayed by sequencing and allele-specific polymerase chain reaction. Polymorphisms in the other odorant receptors in the same odorant receptor gene cluster as *OR7D4* were assayed by sequencing only.

Human olfactory psychophysics. All procedures involving human subjects were approved by the Rockefeller University Institutional Review Board. All subjects completed two replicates of the test separated by at least 4 days. Odours were presented in bar-coded amber vials to ensure that subjects were blind to the identity of all odours²⁸. The intensity and valence of 66 odours at two concentrations ('high' and 'low') and two solvents was rated on a seven-point scale. Thresholds were calculated by using the single-staircase method with seven reversals^{14,15}. Threshold tests included both steroids as binary dilutions from 1:64 (binary dilution 6) to 1:134,217,728 (binary dilution 27). Odour profiling used a previously established method¹⁶.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 June; accepted 8 August 2007.

Published online 16 September 2007.

- Wysocki, C. J. & Beauchamp, G. K. Ability to smell androstenone is genetically determined. *Proc. Natl Acad. Sci. USA* **81**, 4899–4902 (1984).
- Gower, D. B., Nixon, A. & Mallet, A. I. In *Perfumery* (eds Van Toller, S. & Dodd, G. H.) 47–75 (Chapman & Hall, London, 1998).
- Bremner, E. A., Mainland, J. D., Khan, R. M. & Sobel, N. The prevalence of androstenone anosmia. *Chem. Senses* **28**, 423–432 (2003).
- Guillot, M. Physiologie des sensations—anosmies partielles et odeurs fondamentales. *C.R. Hebd. Acad. Sci.* **226**, 1307–1309 (1948).
- Gilbert, A. N. & Kemp, S. E. Odor perception phenotypes: multiple, specific hyperosmias to musks. *Chem. Senses* **21**, 411–416 (1996).
- Whissell-Buechy, D. & Amoore, J. E. Odour-blindness to musk: simple recessive inheritance. *Nature* **242**, 271–273 (1973).
- Menashe, I., Man, O., Lancet, D. & Gilad, Y. Different noses for different people. *Nature Genet.* **34**, 143–144 (2003).
- Gilad, Y. & Lancet, D. Population differences in the human functional olfactory repertoire. *Mol. Biol. Evol.* **20**, 307–314 (2003).
- Saito, H., Kubota, M., Roberts, R. W., Chi, Q. & Matsunami, H. RTP family members induce functional expression of mammalian odorant receptors. *Cell* **119**, 679–691 (2004).
- Keller, A. & Vosshall, L. B. Human olfactory psychophysics. *Curr. Biol.* **14**, R875–R878 (2004).
- Stevens, D. A. & O'Connell, R. J. Enhanced sensitivity to androstenone following regular exposure to pemenone. *Chem. Senses* **20**, 413–419 (1995).
- Zhuang, H. & Matsunami, H. Synergism of accessory factors in functional expression of mammalian odorant receptors. *J. Biol. Chem.* **282**, 15284–15293 (2007).
- Zhang, X. *et al.* Characterizing the expression of the human olfactory receptor gene family using a novel DNA microarray. *Genome Biol.* **8**, R86 (2007).
- Doty, R. L., McKeown, D. A., Lee, W. W. & Shaman, P. A study of the test–retest reliability of ten olfactory tests. *Chem. Senses* **20**, 645–656 (1995).
- Doty, R. L. & Laing, D. G. in *Handbook of Olfaction and Gustation* (ed. Doty, R. L.) 203–228 (Marcel Dekker, New York, 2003).
- Dravnieks, A. Odor quality: semantically generated multidimensional profiles are stable. *Science* **218**, 799–801 (1982).
- Cleveland, W. S. Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* **74**, 829–836 (1979).
- Malnic, B., Hirono, J., Sato, T. & Buck, L. B. Combinatorial receptor codes for odors. *Cell* **96**, 713–723 (1999).
- Wysocki, C. J., Dorries, K. M. & Beauchamp, G. K. Ability to perceive androstenone can be acquired by ostensibly anosmic people. *Proc. Natl Acad. Sci. USA* **86**, 7976–7978 (1989).
- Dorries, K. M., Schmidt, H. J., Beauchamp, G. K. & Wysocki, C. J. Changes in sensitivity to the odor of androstenone during adolescence. *Dev. Psychobiol.* **22**, 423–435 (1989).
- Dalton, P., Doolittle, N. & Breslin, P. A. Gender-specific induction of enhanced sensitivity to odors. *Nature Neurosci.* **5**, 199–200 (2002).
- Mainland, J. D. *et al.* One nostril knows what the other learns. *Nature* **419**, 802 (2002).
- Wang, L., Chen, L. & Jacob, T. Evidence for peripheral plasticity in human odour response. *J. Physiol. (Lond.)* **554**, 236–244 (2004).
- Jacob, S., Kinnunen, L. H., Metz, J., Cooper, M. & McClintock, M. K. Sustained human chemosignal unconsciously alters brain function. *Neuroreport* **12**, 2391–2394 (2001).
- Wyart, C. *et al.* Smelling a single component of male sweat alters levels of cortisol in women. *J. Neurosci.* **27**, 1261–1265 (2007).
- Krautwurst, D., Yau, K. W. & Reed, R. R. Identification of ligands for olfactory receptors by functional expression of a receptor library. *Cell* **95**, 917–926 (1998).
- Laird, D. W. & Molday, R. S. Evidence against the role of rhodopsin in rod outer segment binding to RPE cells. *Invest. Ophthalmol. Vis. Sci.* **29**, 419–428 (1988).
- Keller, A. & Vosshall, L. B. A psychophysical test of the vibration theory of olfaction. *Nature Neurosci.* **7**, 337–338 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements L.B.V. and A.K. thank E. Gotschlich, B. Coller, A. N. Gilbert, I. Gomez, P. Hempstead and C. Vancil; H.M. and H.Z. thank H. Amrein, M. Cook, M. Kubota, D. Marchuk, R. Molday, D. Tracey and R. Valdivia. This research was supported in part by an NIH Clinical and Translational Science Award to Rockefeller University and by grants to L.B.V. from the Irma T. Hirsch Trust, to H.M. from the NIH, to H.Z. from an NIH National Research Service Award, and to A.K. from a Marco S. Stoffel Fellowship.

Author Contributions H.Z. and H.M. screened for androstenone receptors, identified polymorphisms, performed functional expression of receptor variants, and genotyped the human subjects with assistance from Q.C. A.K. and L.B.V. devised the human olfactory psychophysics study, for which A.K. supervised data collection and analysis.

Author Information The sequences of the human *OR7D4* variants are deposited in Genbank under accession numbers EU049291–EU049294. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the paper on www.nature.com/nature. Correspondence and requests for materials should be addressed to H.M. (hiroaki.matsunami@duke.edu) and L.B.V. (leslie@mail.rockefeller.edu).

METHODS

Heterologous expression of human odorant receptors. For untagged odorant receptor experiments, OR7D4 RT and S84N variants without the Rho tag were cloned into pCI. Luciferase assays were performed as described⁹. Only a single variant of each receptor was used in the functional screen, and it is known that many human odorant receptor genes are highly polymorphic⁷. Because of this and because we do not know whether each odorant receptor is capable of functional expression in the cell line used, we cannot exclude the possibility that additional human odorant receptors respond to androstenone as strongly as OR7D4. For immunocytochemistry, cells were fixed, permeabilized and incubated with monoclonal anti-rhodopsin antibody, 4D2 (ref. 27), followed by Cy3-conjugated donkey anti-mouse IgG (Jackson Immunologicals). Western blot analysis was performed in accordance with the Mini-Protein 2 Cell (Bio-Rad) protocol. Enhanced chemiluminescence (ECL; Amersham) was used for detecting proteins on membranes. After the initial exposure, the membrane was incubated with stripping buffer (25 mM glycine-HCl pH 2, 1% SDS, 25 mM glycine, 0.036 M HCl, 1% SDS) and incubated with rabbit anti-GFP (Invitrogen). See Supplementary Methods for detailed information.

Odours. All odours were supplied by Sigma-Aldrich, with these exceptions: androstadienone (gift from Human Pheromone Sciences, Inc.), banana (Bell Flavors and Fragrances), bourgeonal (Biomol), galaxolide (gift from International Flavors and Fragrances) and (*R*)-carvone (Research Chemical Ltd). The same batch and lot of each odour was used for both cell-based analysis and human olfactory psychophysics. Detailed information on odours, odour concentrations and perceived odour quality is provided in Supplementary Tables 4–6.

Genotyping and sequencing of human odorant receptors. For sequencing, human genomic DNAs were amplified, purified and sequenced with a 3100 or 3730 Genetic Analyser (ABI Biosystems) or by GeneWiz. Detailed methods are given in Supplementary Methods.

Human olfactory psychophysics. All human subjects gave informed consent to participate in this study and were tested in a well-ventilated room of the Rockefeller University Hospital Outpatient Unit. Normal human subjects were pre-screened to exclude pregnant women and those with medical conditions causing general impairment of the sense of smell. Of the 412 subjects who completed the study, 21 were excluded because of general anosmia (see Supplementary Methods). The remaining 391 subjects (210 female, 181 male; median age 34 years, age range 19–75 years) were included in the evaluation. Detailed methods are given in Supplementary Information. Our smell tests were purposely conducted under conditions that would not be expected to induce odour sensitivity in our subjects. A given subject sniffed androstenone in only two sessions, rather than the dozens of sessions spread over six to eight weeks required for sensitization in previous studies¹⁹. A comparative analysis of androstenone responses in the first and second visits does not suggest that subjects became more sensitive through these brief, non-chronic exposures (data not shown).

Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors

César Cobaleda^{1†}, Wolfram Jochum² & Meinrad Busslinger¹

Lineage commitment and differentiation to a mature cell type are considered to be unidirectional and irreversible processes under physiological conditions¹. The commitment of haematopoietic progenitors to the B-cell lineage^{2,3} and their development to mature B lymphocytes^{4,5} critically depend on the transcription factor encoded by the paired box gene 5 (*Pax5*). Here we show that conditional *Pax5* deletion in mice allowed mature B cells from peripheral lymphoid organs to dedifferentiate *in vivo* back to early uncommitted progenitors in the bone marrow, which rescued T lymphopoiesis in the thymus of T-cell-deficient mice. These B-cell-derived T lymphocytes carried not only immunoglobulin heavy- and light-chain gene rearrangements but also participated as functional T cells in immune reactions. Mice lacking *Pax5* in mature B cells also developed aggressive lymphomas, which were identified by their gene expression profile as progenitor cell tumours. Hence, the complete loss of *Pax5* in late B cells could initiate lymphoma development and uncovered an extraordinary plasticity of mature peripheral B cells despite their advanced differentiation stage.

The B-cell-specific *Cd19-cre* gene⁶ efficiently deletes the floxed (*fl*) *Pax5* allele only in B cells of peripheral lymphoid organs, which facilitated the analysis of *Pax5* function in mature B cells of young *Cd19-cre Pax5^{fl/fl}* mice⁴. In an ageing mouse colony, half of the *Cd19-cre Pax5^{fl/fl}* mice died within eight months, unlike the *Cd19-cre Pax5^{fl/+}* mice retaining a functional *Pax5* allele (Fig. 1a). All diseased *Cd19-cre Pax5^{fl/fl}* mice developed aggressive lymphoma resulting in splenomegaly and massively enlarged lymph nodes (Fig. 1b, c) owing to diffuse infiltration by lymphoma cells, which were also found in many other organs (Supplementary Table 1), including the lung, liver, kidney and thymus (Fig. 1d–g). The lymphoma cells could readily be established in culture and rapidly developed into secondary lymphomas after transplantation into syngeneic mice (Supplementary Table 1). Spectral karyotype analysis identified two distinct chromosomal translocations in two lymphomas but a normal diploid karyotype in seven additional tumours, indicating that a specific translocation or high genomic instability is not required for lymphoma development in *Cd19-cre Pax5^{fl/fl}* mice (Supplementary Fig. 1). *Cd19-cre Pax5^{fl/fl} Eμ-bcl2* mice died after a shorter latency period, demonstrating that Bcl2-mediated cell survival cooperates with *Pax5* loss in lymphoma development (Fig. 1a).

Flow cytometric analysis revealed large blastic *Cd19-cre Pax5^{fl/fl}* lymphoma cells compared to small quiescent B cells in control *Cd19-cre Pax5^{fl/+}* lymph nodes (Fig. 1h). The tumour cells failed to express the B-cell surface proteins CD19, CD21, CD22, CD23, CD40, CD72, immunoglobulin (Ig)D and major histocompatibility complex class II (Fig. 1h and Supplementary Table 2), all of which are also downregulated on conditional *Pax5* loss in mature B cells⁴. Consistent with a B-lymphoid origin, all lymphoma cells expressed

B220 and the immunoglobulin Igμ protein (which was part of the pre-BCR) in the absence of Igκ and Igλ expression (Fig. 1h and Supplementary Table 2). In addition, the tumour cells expressed the early markers CD93 and CD25 together with the signalling receptors interleukin (IL)-7Rα, c-Kit and Flt3, which are normally present on lymphocyte progenitors (Fig. 1h). Because *Flt3* is a repressed *Pax5* target gene^{7,8}, its expression indicated that the *Cd19-cre Pax5^{fl/fl}* lymphomas may correspond to *Pax5*-deficient progenitor cell tumours. Complementary DNA microarray analysis together with *Pax5* genotyping confirmed this hypothesis, because the lymphoma cells were indistinguishable from *Pax5^{-/-}* pro-B cells with respect to the transcription of *Pax5*-repressed and *Pax5*-activated genes (Supplementary Fig. 2a, b).

To investigate the stage-specific origin of the *Cd19-cre Pax5^{fl/fl}* tumours, we analysed their *Igh*, *Igk* and *Igl* rearrangements by polymerase chain reaction (PCR) (Fig. 1i), Southern blotting (Supplementary Fig. 3), sequencing of individual PCR fragments (Supplementary Fig. 2c) and PCR with reverse transcription (RT-PCR) of rearranged immunoglobulin transcripts (Supplementary Fig. 4). All tumours contained a single, in-frame rearranged and expressed *Igh* allele (predominantly containing a *VHJ558*, also known as *Igh-VJ558*, gene), indicating that the tumours were of clonal origin. Most tumours also carried an in-frame rearranged *Igk* or *Igl* allele. It is unlikely that these rearrangements arose in pro-B cells undergoing *Pax5* deletion, because *Pax5*-deficient pro-B cells are unable to generate *VHJ558-DJH* and *Vκ-Jκ* rearrangements^{9,10}. Instead, late pre-B cells correspond to the first developmental stage carrying functional immunoglobulin heavy- and light-chain gene rearrangements. We therefore conclude that the progenitor cell lymphomas in *Cd19-cre Pax5^{fl/fl}* mice must originate at least from late pre-B cells or, more probably, from immature or mature B cells.

To demonstrate developmental plasticity of mature B cells in the absence of oncogenic events, we used an experimental strategy that relied on the isolation of highly purified *Pax5*-deleted mature B cells. *Pax5* deletion was achieved either *in vivo* with *Cd19-cre* or *in vitro* with the *CreED-30* transgene encoding a hormone-inducible Cre-oestrogen receptor fusion (Cre-ER) protein¹¹. Before the isolation of mature B cells, the *CreED-30 Pax5^{fl/fl} Eμ-bcl2 Ly5.2⁺* mice (used for *in vitro Pax5* deletion) and *Cd19-cre Pax5^{fl/fl} Eμ-bcl2 Ly5.2⁺* mice (used for *in vivo Pax5* deletion) were repeatedly injected with anti-IL-7Rα antibody to block B-cell development¹², thus eliminating contaminating precursor and immature B cells in peripheral lymphoid organs (Supplementary Fig. 5). Subsequently, Lin[−]IgM⁺IgD^{high} mature B cells were FACS-sorted with a purity of >99% from the spleen or lymph nodes of *CreED-30 Pax5^{fl/fl} Eμ-bcl2 Ly5.2⁺* mice (Fig. 2a and Supplementary Fig. 6a). The isolated B cells were treated *in vitro* for 40 h with 4-hydroxytamoxifen (OHT) to delete the floxed *Pax5* allele before injection into *Rag2^{-/-} Ly5.1⁺* mice (Fig. 2a). *In vivo*

¹Research Institute of Molecular Pathology, Vienna Biocenter, Dr. Bohr-Gasse 7, A-1030 Vienna, Austria. ²Department of Pathology, University Hospital, Schmelzbergstrasse 12, CH-8091 Zürich, Switzerland. †Present address: Universidad de Salamanca, Campus Miguel de Unamuno, 37007 Salamanca, Spain.

Pax5-deleted mature B cells were purified as $\text{Lin}^- \text{CD}25^+ \text{IgM}^+ \text{IgD}^-$ cells⁴ from the lymph nodes of *Cd19-cre Pax5^{fl/fl} Eμ-bcl2 Ly5.2⁺* mice (Supplementary Fig. 6b) followed by transplantation into *Rag2^{-/-} Ly5.1⁺* mice (Fig. 2a). The allogeneic markers *Ly5.1* and *Ly5.2* were

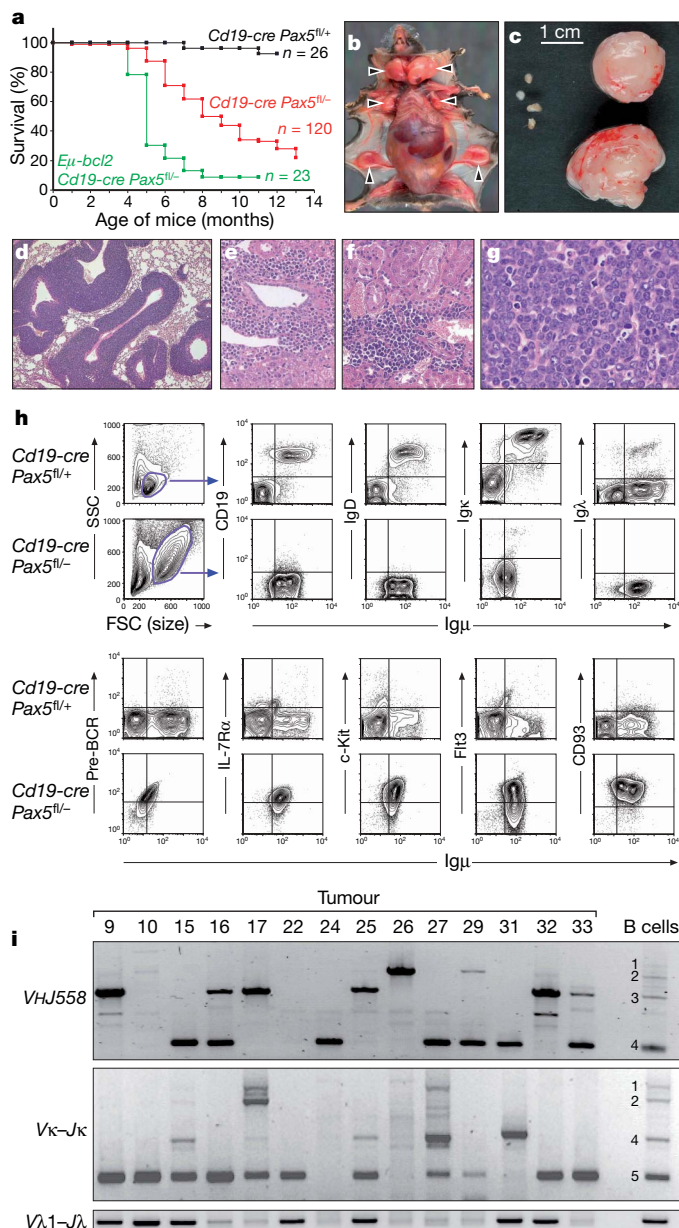


Figure 1 | Development of aggressive progenitor cell lymphomas on *Pax5* loss in B lymphocytes. **a**, Kaplan-Meier survival analysis of mice of the indicated genotypes. The number (*n*) of mice analysed is shown. **b**, Enlarged lymph nodes (arrowheads) in a *Cd19-cre Pax5^{fl/fl}* tumour mouse. **c**, Size comparison of lymph nodes from a *Cd19-cre Pax5^{fl/fl}* tumour mouse (right) and a control littermate (left). **d–g**, Eosin-haematoxylin-stained sections of the lung (**d**), liver (**e**), kidney (**f**) and thymus (**g**) of a *Cd19-cre Pax5^{fl/fl}* tumour mouse. The infiltration patterns of all tumour mice are shown in Supplementary Table 1. **h**, Flow cytometric analysis of lymph node cells of a *Cd19-cre Pax5^{fl/fl}* tumour mouse and a control *Cd19-cre Pax5^{fl/fl}* littermate. Cells within the forward and side scatter (FSC/SSC) gates (blue) were analysed. The entire immunophenotype of *Cd19-cre Pax5^{fl/fl}* tumour cells is summarized in Supplementary Table 2. **i**, PCR analysis of *VHJ558-DJH*, *Vκ-Jκ* and *Vλ1-Jλ* rearrangements in tumour-infiltrated lymph nodes of *Cd19-cre Pax5^{fl/fl}* mice. Rearrangements to all *JH* and *Jκ* segments (numbered) were detected in splenic B cells. The detailed characterization of the different immunoglobulin gene rearrangements by cloning and sequencing is summarized in Supplementary Fig. 2c.

used to discriminate haematopoietic cells of host (*Ly5.1⁺*) and donor (*Ly5.2⁺*) origin in the reconstituted mice.

Eight weeks after transplantation, *Ly5.2⁺ c-Kit⁺ B220⁺* pro-B cells of donor origin could be detected in the bone marrow of the 19 mice analysed, regardless of whether these mice were injected with *in vitro* or *in vivo* *Pax5*-deleted mature B cells (Fig. 2b). These *Ly5.2⁺ c-Kit⁺ B220⁺* pro-B cells were identified as *Pax5*-deficient progenitors, because they did not express the *Pax5* target gene *Cd19* (ref. 13) (Fig. 2b). Mature splenic B cells purified from control *Pax5^{fl/fl} Eμ-bcl2 Ly5.2⁺* mice (lacking *CreED-30*) or *CreED-30 Pax5^{fl/fl} Ly5.2⁺* mice (lacking *Eμ-bcl2*) failed to give rise to *Ly5.2⁺* pro-B cells in seven and four injected *Rag2^{-/-} Ly5.1⁺* recipients, respectively (data not shown). Hence, mature B cells require *Pax5* inactivation as well as prolonged cell survival to be able to dedifferentiate to progenitor cells *in vivo*. The requirement of transgenic *Bcl2* expression is consistent with a slow dedifferentiation process, because *in vitro* *Pax5*-deleted B cells labelled with carboxyfluorescein succinimidyl ester could still be detected as *CD19⁺ IgM⁺ IgD⁺ c-Kit⁻* B lymphocytes in the spleen of *Rag2^{-/-} Ly5.1⁺* mice one week after injection (Supplementary Fig. 7).

The *Ly5.2⁺ c-Kit⁺ B220⁺ CD19⁻* pro-B cells isolated from the bone marrow of reconstituted mice could be cultured under lymphoid conditions as uncommitted progenitors that were able to differentiate *in vitro* into macrophages after IL-7 withdrawal and addition of macrophage-colony-stimulating factor (Supplementary Fig. 8a–c). Eight weeks after transplantation of *Pax5*-deleted B cells, myeloid *Ly5.2⁺ Mac1⁺ Gr1⁻* cells of donor origin could also be detected in the spleen of four of the 12 mice analysed (Supplementary Fig. 8d), in agreement with the inefficient *in vivo* myeloid differentiation of *Pax5^{-/-}* pro-B cells¹⁴.

In contrast, all *Rag2^{-/-} Ly5.1⁺* mice showed robust reconstitution of T-cell development after transplantation with *in vitro* or *in vivo* *Pax5*-deleted B cells (Fig. 2b and Supplementary Fig. 9), consistent with the fact that *Pax5*-deficient progenitors from the bone marrow efficiently seed the thymus to initiate T lymphopoiesis³. Within 2–3 months, donor-derived *Ly5.2⁺* T cells restored the thymic cellularity to one-third of that of wild-type mice (Fig. 3c), revealed a normal distribution of *CD4⁺ CD8⁻* (double-negative), *CD4⁺ CD8⁺* (double-positive), *CD4⁺* (single-positive) and *CD8⁺* (single-positive) T cells (Fig. 2b and Supplementary Fig. 9) and carried polyclonal T-cell receptor- β gene (*Tcrb*) rearrangements like normal T cells (Fig. 3a, b and Supplementary Fig. 10). PCR analysis of *Igh*, *Igk* and *Igl* rearrangements confirmed the B-cell origin of the *Ly5.2⁺* thymocytes (Fig. 3a, b). In contrast to wild-type T cells, the sorted *Ly5.2⁺* double-positive thymocytes carried oligoclonal *VH-DJH* rearrangements of distinct *Igh* *VH* gene families in addition to *Vκ-Jκ* and *Vλ1-Jλ* rearrangements (Fig. 3a, b and Supplementary Table 3).

PCR analysis revealed a similar oligoclonality of the immunoglobulin gene rearrangements in *Ly5.2⁺* pro-B cells from the bone marrow of reconstituted mice (Supplementary Fig. 11), which further identified the dedifferentiation of *Pax5*-deleted mature B cells to uncommitted progenitors as the rate-limiting step. Serial transplantation experiments demonstrated that the *Ly5.2⁺* pro-B cells isolated from a primary reconstituted *Rag2^{-/-} Ly5.1⁺* mouse homed back to the bone marrow and reconstituted T-cell development in the thymus and spleen after injection into secondary *Rag2^{-/-} cγ^{-/-}* recipients (Supplementary Fig. 12). These serial transplantation data and the normal developmental progression of T lymphopoiesis in reconstituted *Rag2^{-/-}* mice rule out transdifferentiation as a mechanism for the conversion of mature B cells into T cells. Instead, the loss of *Pax5* allows mature B cells to develop into T cells in the thymus by dedifferentiation to an uncommitted progenitor cell stage in the bone marrow.

Double-positive thymocytes of reconstituted *Rag2^{-/-}* mice exhibited a normal proliferation response and cytokine production on appropriate stimulation (Supplementary Fig. 13). To study the *in vivo* immune response of B-cell-derived T cells, we used *Tcra^{-/-}* mice¹⁵ as recipients for the injection of *in vitro* *Pax5*-deleted mature

B cells (Fig. 4a). Eight weeks after transplantation, the reconstituted *Tcra*^{-/-} mice were immunized with sheep red blood cells, and the germinal-centre reaction was investigated by flow cytometric and immunohistochemical analyses of the spleen ten days after immunization (Fig. 4b, c). In contrast to the lymphopenic *Rag2*^{-/-} mice, the *Tcra*^{-/-} mice have normal B-cell compartments and an almost normal thymic cellularity, because they only fail to develop CD4 and CD8 single-positive T cells¹⁵. Owing to the absence of TCR-β⁺ T cells in the spleen, the *Tcra*^{-/-} mice are unable to generate germinal centres and PNA⁺Fas⁺ germinal-centre B cells in response to immunization (Fig. 4b, c). In contrast, the spleen of the reconstituted *Tcra*^{-/-} mice contained polyclonal TCR-β⁺ T cells (Fig. 4b) carrying wild-type *Tcra* and deleted *Pax5* alleles (Supplementary Fig. 14). These donor-derived T cells efficiently induced the development of germinal centres and PNA⁺Fas⁺ germinal-centre B cells (Fig. 4b, c), indicating that the B-cell-derived T cells are fully functional.

Ectopic expression of lineage-specific transcription factors can convert one cell type into another, as documented by the *in vitro* transdifferentiation of B lymphocytes into macrophages in response to retroviral expression of the CCAAT/enhancer-binding protein-α (ref. 16). Moreover, forced expression of the transcription factors Oct3/4, Sox2, c-Myc and Klf4 is sufficient to transdifferentiate mouse fibroblasts *in vitro* into germline-competent embryonic stem cells^{17,18}. In contrast to these gain-of-function experiments, we have demonstrated that the loss of a single transcription factor, Pax5, allows mature B cells to dedifferentiate *in vivo* to uncommitted haematopoietic progenitors and to develop into progenitor cell lymphomas. *Pax5* inactivation also induces the dedifferentiation of committed pro-B cells, which have entered the B-cell lineage for only one

or two days¹⁹. In marked contrast, the mature B cells passed several B-cell developmental checkpoints, migrated to peripheral lymphoid organs, are quiescent and, owing to their longevity, are three to five months old^{20,21}. Despite their advanced differentiation stage, mature B cells retain an extraordinary developmental plasticity, as revealed by their dedifferentiation potential on Pax5 loss. Under physiological conditions, *Pax5* and its transcription programme are downregulated during terminal differentiation of mature B cells to plasma cells^{7,22}, which is initiated in response to antigen stimulation of the BCR²³. Hence, the loss of Pax5 in the context of strong BCR signalling results in forward differentiation of mature B cells to plasma cells, whereas *Pax5* inactivation in the absence of BCR signalling initiates the reversal of differentiation to uncommitted progenitors.

Chromosomal translocations implicated PAX5 as an oncogene in the generation of human acute lymphoblastic leukaemias and non-Hodgkin's lymphomas⁵. A genome-wide analysis revealed mono-allelic loss or point mutations of PAX5 in 32% of B-progenitor acute lymphoblastic leukaemias²⁴. These genetic lesions result in haploinsufficiency, suggesting that heterozygous PAX5 mutations contribute to leukaemogenesis²⁴. Our results indicate, however, that the inactivation of one *Pax5* allele in the absence of other oncogenic lesions is not sufficient to induce tumour development in heterozygous *Cd19-cre Pax5*^{fl/+} mice. Instead, the complete loss of Pax5 in B cells leads to an aggressive progenitor cell lymphoma. The Pax5-deficient progenitor cell state probably contributes to lymphomagenesis, because *Pax5*^{-/-} pro-B cells retain an extensive *in vivo* self-renewal potential²⁵ similar to stem cells¹ and cancer stem cells²⁶. Our loss-of-function experiments thus identified Pax5 as a tumour suppressor gene of the

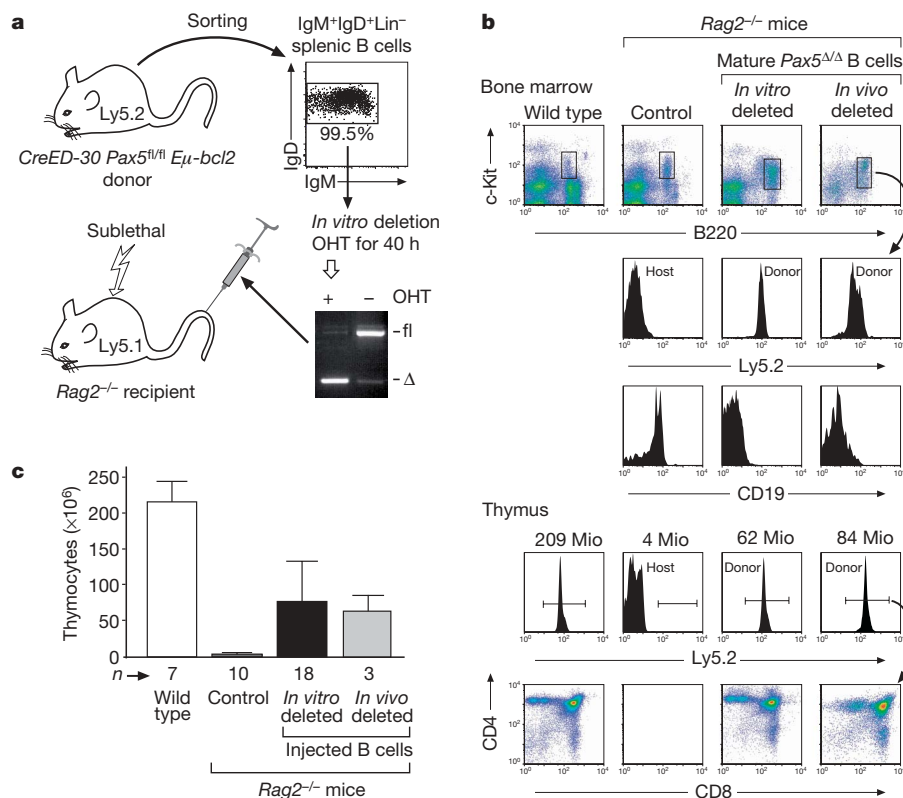


Figure 2 | T-cell reconstitution by Pax5-deleted mature B cells in *Rag2*^{-/-} mice. **a**, Experimental strategy. IgM⁺IgD⁺Lin⁻ B cells from the spleen or lymph nodes of *CreED-30 Pax5*^{fl/fl} *Eμ-bcl2* *Ly5.2*⁺ mice were purified by FACS sorting (reanalysis shown) and incubated *in vitro* with OHT for 40 h before injection into sublethally irradiated *Rag2*^{-/-} *Ly5.1*⁺ mice. Deletion (Δ) of the floxed (fl) *Pax5* allele is shown. **b**, Flow cytometric analysis of the bone marrow and thymus of *Rag2*^{-/-} *Ly5.1*⁺ recipient mice eight weeks after transplantation of *in vitro* or *in vivo* Pax5-deleted mature B cells. *Ly5.2* and

CD19 expression is shown for the gated B220⁺c-Kit⁺ pro-B cells in the bone marrow (upper panels). The thymic cellularity and *Ly5.2* expression of all thymocytes is shown together with the CD4 and CD8 expression pattern of gated *Ly5.2*⁺ thymocytes (lower panels). The analysis of all reconstituted mice is shown in Supplementary Fig. 9. Mio, millions of cells. **c**, Restoration of thymic cellularity in reconstituted *Rag2*^{-/-} mice. The average (+ s.d.) cellularity and number (n) of mice analysed are indicated.

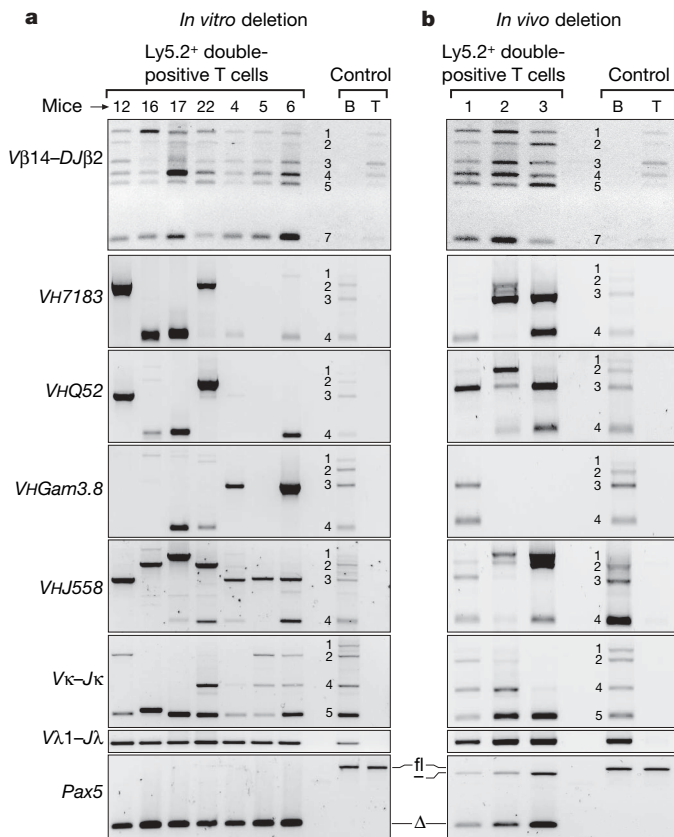


Figure 3 | Immunoglobulin gene rearrangements in thymocytes of reconstituted *Rag2*^{-/-} mice. PCR detection of immunoglobulin gene rearrangements in Ly5.2⁺ CD4⁺ CD8⁺ (double-positive) thymocytes, which were sorted 8–12 weeks after transplantation of *Rag2*^{-/-} Ly5.1⁺ recipient mice with *in vitro* Pax5-deleted mature B cells (a) or *in vivo* Pax5-deleted mature B cells (b). Oligoclonal VH–DJH, Vκ–Jκ and Vλ1–Jλ rearrangements were detected in Ly5.2⁺ double-positive thymocytes in addition to polyclonal Vβ14–Jβ2 rearrangements. B220⁺ splenocytes (B) and double-positive thymocytes (T) from Pax5^{fl/fl} mice were used as controls. Numbers indicate rearrangements involving different Jβ2, JH or Jκ segments. PCR genotyping revealed full conversion of the fl to the Δ Pax5 allele in all Ly5.2⁺ thymocytes. The Pax5 mutant (–) germline allele is indicated. The VHGam3.8, VHQ52 and VH7183 genes are also known as Igh–VGAM3.8, Igh–VQ52 and Igh–V7183, respectively.

B-lymphoid lineage. Hence, Pax5 can function as a tumour suppressor or oncoprotein in the generation of different lymphoid malignancies in analogy to a similar dual role of Notch1 (refs 27, 28) and c-Fos^{29,30} in the development of distinct cancers.

METHODS SUMMARY

Mice. All mouse strains (see Methods) were maintained on the C57BL/6 background.

Tumour analysis. Organs were fixed in 4% formaldehyde followed by paraffin embedding, sectioning and staining with haematoxylin and eosin. Lymph node tumour cells were cultured with γ-irradiated stromal (ST2) cells, IL-7, Flt3L and stem cell factor before injection into syngeneic mice.

***In vitro* Pax5-deleted mature B cells.** CreED-30 Pax5^{fl/fl} Eμ-bcl2 Ly5.2⁺ mice were injected 5–7 times with the anti-IL-7Rα antibody A7R34, and Lin⁺ non-B cells were removed from splenocytes or lymph node cells by magnetic cell sorting (MACS) after staining with lineage marker antibodies (CD3ε, CD4, CD8a, CD11c, CD49b, CD93, Gr1, c-Kit, Mac1, TCR-β, Ter119 and Thy1.2). After FACS sorting, the Lin⁺ IgM⁺ IgD^{high} B cells were incubated *in vitro* with the oestrogen analogue OHT to delete the floxed Pax5 alleles by the OHT-induced Cre-ER activity (CreED-30).

***In vivo* Pax5-deleted mature B cells.** Lymph-node cells from anti-IL-7Rα antibody-treated Cd19-cre Pax5^{fl/fl} Eμ-bcl2 Ly5.2⁺ mice were depleted of Lin⁺ non-B cells with a similar lineage antibody cocktail before sorting of Pax5-deficient mature B cells as Lin⁺ CD25⁺ IgM⁺ IgD⁺ cells⁴.

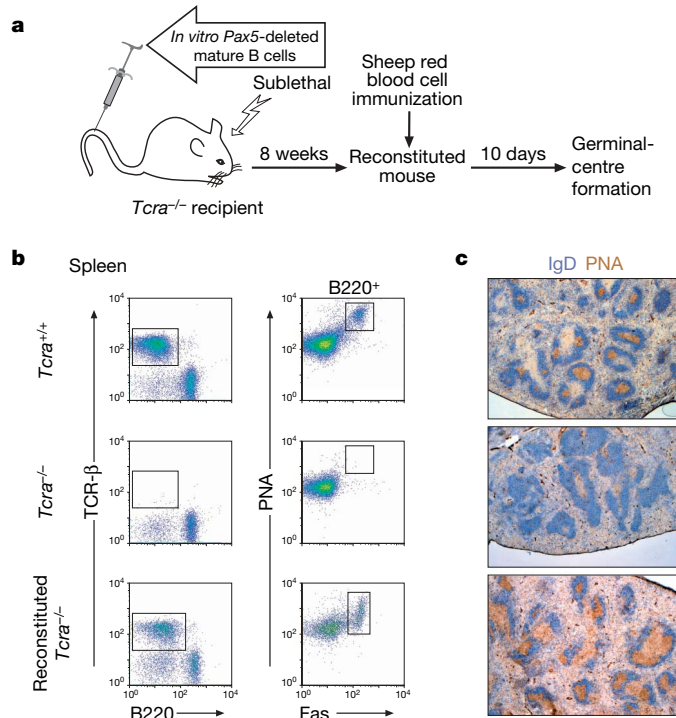


Figure 4 | Normal immune function of B-cell-derived T cells. a, Schematic diagram of the reconstitution and immunization of *Tcrα*^{-/-} mice with sheep red blood cells. b, Flow cytometric analysis of splenocytes from mice of the indicated genotypes ten days after immunization with sheep red blood cells. c, Formation of germinal centres in the same immunized mice shown in b. Cryosections of the spleen were stained with fluorescein isothiocyanate (FITC)–anti-IgD antibody (blue) and biotinylated peanut haemagglutinin (PNA, brown) followed by detection with an alkaline phosphatase-coupled anti-FITC antibody (visualized with Fast Blue) and horseradish peroxidase-conjugated streptavidin (visualized with diaminobenzidine), respectively.

B-cell transplantations. *In vitro* Pax5-deleted B cells (2–4 × 10⁶ cells per mouse) or *in vivo* Pax5-deleted B cells (5 × 10⁵ cells per mouse) were intravenously injected into 8–12-week-old *Rag2*^{-/-} Ly5.1⁺ or *Tcrα*^{-/-} recipient mice that were γ-irradiated with a sublethal dose (4 Gy).

V(D)J recombination. Igh, Igk, Igl and Tcrb rearrangements were amplified by PCR using the primers listed in Supplementary Table 4. To characterize individual V–(D)J rearrangements, the PCR fragments were cloned, and at least ten clones corresponding to the same PCR fragment were sequenced.

Immunizations. Sheep red blood cells (1–2 × 10⁸ cells) were intraperitoneally injected into control or reconstituted *Tcrα*^{-/-} mice. Ten days later, the spleen was analysed by flow cytometry and immunohistochemistry⁷.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 2 May; accepted 8 August 2007.

Published online 12 September 2007.

- Weissman, I. L. Stem cells: units of development, units of regeneration, and units of evolution. *Cell* **100**, 157–168 (2000).
- Nutt, S. L., Heavey, B., Rolink, A. G. & Busslinger, M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401**, 556–562 (1999).
- Rolink, A. G., Nutt, S. L., Melchers, F. & Busslinger, M. Long-term *in vivo* reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature* **401**, 603–606 (1999).
- Horchner, M., Souabni, A. & Busslinger, M. Pax5/BSAP maintains the identity of B cells in late B lymphopoiesis. *Immunity* **14**, 779–790 (2001).
- Cobaleda, C., Schebesta, A., Delogu, A. & Busslinger, M. Pax5: the guardian of B cell identity and function. *Nature Immunol.* **8**, 463–470 (2007).
- Rickert, R. C., Roes, J. & Rajewsky, K. B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic Acids Res.* **25**, 1317–1318 (1997).
- Delogu, A. *et al.* Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity* **24**, 269–281 (2006).
- Holmes, M. L., Carotta, S., Corcoran, L. M. & Nutt, S. L. Repression of *Flt3* by Pax5 is crucial for B-cell lineage commitment. *Genes Dev.* **20**, 933–938 (2006).

9. Fuxa, M. *et al.* Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev.* **18**, 411–422 (2004).
10. Sato, H., Saito-Obara, F., Inazawa, J. & Kudo, A. Pax-5 is essential for κ sterile transcription during Ig κ chain gene rearrangement. *J. Immunol.* **172**, 4858–4865 (2004).
11. Schwenk, F., Kühn, R., Angrand, P.-O., Rajewsky, K. & Stewart, A. F. Temporally and spatially regulated somatic mutagenesis in mice. *Nucleic Acids Res.* **26**, 1427–1432 (1998).
12. Sudo, T. *et al.* Expression and function of the interleukin 7 receptor in murine lymphocytes. *Proc. Natl Acad. Sci. USA* **90**, 9125–9129 (1993).
13. Nutt, S. L., Morrison, A. M., Dörfler, P., Rolink, A. & Busslinger, M. Identification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments. *EMBO J.* **17**, 2319–2333 (1998).
14. Schaniel, C., Bruno, L., Melchers, F. & Rolink, A. G. Multiple hematopoietic cell lineages develop *in vivo* from transplanted Pax5-deficient pre-B I-cell clones. *Blood* **99**, 472–478 (2002).
15. Philpott, K. L. *et al.* Lymphoid development in mice congenitally lacking T cell receptor $\alpha\beta$ -expressing cells. *Science* **256**, 1448–1452 (1992).
16. Xie, H., Ye, M., Feng, R. & Graf, T. Stepwise reprogramming of B cells into macrophages. *Cell* **117**, 663–676 (2004).
17. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
18. Wernig, M. *et al.* *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318–324 (2007).
19. Mikkola, I., Heavey, B., Horcher, M. & Busslinger, M. Reversion of B cell commitment upon loss of Pax5 expression. *Science* **297**, 110–113 (2002).
20. Förster, I. & Rajewsky, K. The bulk of the peripheral B-cell pool in mice is stable and not rapidly renewed from the bone marrow. *Proc. Natl Acad. Sci. USA* **87**, 4781–4784 (1990).
21. Rolink, A. G., Andersson, J. & Melchers, F. Characterization of immature B cells by a novel monoclonal antibody, by turnover and by mitogen reactivity. *Eur. J. Immunol.* **28**, 3738–3748 (1998).
22. Schebesta, A. *et al.* Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration and immune function. *Immunity* **27**, 49–63 (2007).
23. Shapiro-Shelef, M. & Calame, K. Regulation of plasma-cell development. *Nature Rev. Immunol.* **5**, 230–242 (2005).
24. Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758–764 (2007).
25. Schaniel, C., Gottar, M., Roosnek, E., Melchers, F. & Rolink, A. G. Extensive *in vivo* self-renewal, long-term reconstitution capacity, and hematopoietic multipotency of Pax5-deficient precursor B-cell clones. *Blood* **99**, 2760–2766 (2002).
26. Dalerba, P., Cho, R. W. & Clarke, M. F. Cancer stem cells: models and concepts. *Annu. Rev. Med.* **58**, 267–284 (2007).
27. Nicolas, M. *et al.* Notch1 functions as a tumor suppressor in mouse skin. *Nature Genet.* **33**, 416–421 (2003).
28. Weng, A. P. *et al.* Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* **306**, 269–271 (2004).
29. Grigoriadis, A. E., Schellander, K., Wang, Z.-Q. & Wagner, E. F. Osteoblasts are target cells for transformation in c-fos transgenic mice. *J. Cell Biol.* **122**, 685–701 (1993).
30. Fleischmann, A., Jochum, W., Eferl, R., Witowsky, J. & Wagner, E. F. Rhabdomyosarcoma development in mice lacking Trp53 and Fos: tumor suppression by the Fos protooncogene. *Cancer Cell* **4**, 477–482 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Schebesta for cDNA microarray analysis, M. Murphy and F. Alt for spectral karyotype analysis of tumour cells, F. Alt for providing immunoglobulin gene probes, A. Rolink for suggesting the *Tcra*^{-/-} mouse reconstitution experiment, K. Rajewsky, J. Adams and W. Ellmeier for providing transgenic mouse strains, G. Stengl for FACS sorting, S. Höflinger for help with the V(D)J recombination analysis, and I. Botto and M. Madalinsky for anti-IL-7R α antibody purification. This research was supported by Boehringer Ingelheim (M.B.), the Austrian Industrial Research Promotion Fund (M.B.), a Spanish 'Ramon y Cajal' investigator grant (C.C.), the Fondo de Investigaciones Sanitarias (C.C.), the Junta de Castilla y León (C.C.) and the Fundación de Investigación MMA (C.C.).

Author Contributions C.C. carried out almost all experimental work and contributed to manuscript writing. W.J. performed the histological analyses and evaluation of the tumours. M.B. contributed to the project planning and wrote the manuscript.

Author Information All microarray data have been deposited in the GEO repository at NCBI under accession numbers GSM210098, GSM210099, GSM210100, GSM210101, GSM215734, GSM215735 and GSM213736. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.B. (busslinger@imp.ac.at).

METHODS

Mice. The following mice were maintained on the C57BL/6 background and genotyped as described: *Pax5*^{+/-} (ref. 31), *Pax5*^{fl/fl} (ref. 4), *Rag2*^{-/-} (ref. 32), *Tcrα*^{-/-} (ref. 15), *Cd19*^{cre/cre} (ref. 33), *CreED-30* (ref. 11) and *Em-bcl2-36* (ref. 34). Heterozygous *Cd19*^{cre/+} mice are referred to as *Cd19-cre* mice in this manuscript. All mice were *Ly5.2*⁺ except for the *Rag2*^{-/-} *Ly5.1*⁺ mice, which were obtained from Taconic (000461-M).

FACS sorting and analysis. FITC-, phycoerythrin (PE)- or allophycocyanin (APC)-coupled antibodies to the following cell-surface proteins were used for flow cytometry: cell-surface protein B220 (antibody RA3-6B2), CD3ε (145-2C11), CD4 (H129.19), CD8α (53-6.7), CD11b (also known as Mac1, CD11b/Mac1) (M1/70), CD11c (HL3), CD19 (1D3), CD21 (7G6), CD22 (Cy34.1), CD23 (B3B4), CD25 (PC61), CD40 (FGK45.5 or 3/23), CD49b (DX5), CD72 (K10.6), CD90.2/Thy1.2 (53-2.1), CD93/AA4.1 (PB.493), CD95/Fas (Jo2), CD117/c-Kit (2B8), CD127/IL-7Rα (A7R34), CD135/Flt3 (A2F10.1), CD138 (281-2), F4/80 (CI-A3-1), Gr1 (RB6-8C5), IgD (1.19), Igκ (187.1), Igλ (R26-46), IgM (M41-42), Ly5.1 (A20), Ly5.2 (104), M-CSFR (macrophage-colony-stimulating factor receptor) (AFS98), MHCII (M5-114), pre-BCR (SL156), TCR-β (H57-597) and Ter119 (TER-119). FITC-labelled PNA (Vector Laboratories) was used for staining of germinal-centre B cells. Red blood cells were eliminated from splenocytes by lysis in 150 mM NH₄Cl and 10 mM KHCO₃ for 5 min at 22 °C. Antibody staining and washing of single-cell suspensions were done in phosphate buffered saline solution (PBS) containing 2% heat-inactivated fetal calf serum (FCS). Unspecific antibody binding was suppressed by preincubation of cells with CD16/CD32 Fc-block solution (BD Pharmingen). For all analyses, a forward and side scatter gate was used to include viable cells while excluding dead cells, debris and cell doublets. If possible, dead cells were also eliminated by staining with propidium iodide (1 µg ml⁻¹). A FACS Aria cell sorter (Becton-Dickinson) was used to purify the different B- and T-cell types. **Anti-IL-7Rα antibody treatment.** The purified anti-IL-7Rα antibody A7R34 (1 mg in 150 µl PBS) was injected into the tail vein of *CreED-30 Pax5*^{fl/fl} *Em-bcl2* *Ly5.2*⁺ and *Cd19-cre Pax5*^{fl/+} *Em-bcl2* *Ly5.2*⁺ mice at the age of 6–8 weeks. This injection was repeated every second day during 8–12 days (5–7 injections per mouse), which blocked B-cell development leading to the loss of pro-B, pre-B and immature B cells in the bone marrow and spleen (Supplementary Fig. 5). Mature B cells were sorted two days after the last injection.

In vitro deletion of Pax5 in sorted mature B cells. Single-cell suspensions of the spleen or lymph nodes from anti-IL-7Rα antibody-treated *CreED-30 Pax5*^{fl/fl} *Em-bcl2* *Ly5.2*⁺ mice were stained with PE-coupled lineage marker antibodies (CD3ε, CD4, CD8α, CD11c, CD49b, CD93, Gr1, c-Kit, Mac1, TCR-β, Ter119 and Thy1.2), and the Lin⁺ non-B cells and traces of residual CD93⁺ B-cell precursors were eliminated by MACS with anti-PE beads (Miltenyi Biotec). Mature B cells were then FACS-sorted as Lin⁻ IgM⁺ IgD^{high} cells after staining with FITC-anti-IgM and APC-anti-IgD antibodies. Flow cytometric reanalysis of the sorted mature B cells revealed a purity of more than 99% (Fig. 2a and Supplementary Fig. 6a). The purified B cells were subsequently incubated for 40 h in Iscove's modified Dulbecco's medium supplemented with 10% FCS, 50 µM β-mercaptoethanol, 1 mM glutamine and 1 µM OHT (Sigma). Deletion of the floxed *Pax5* allele by the OHT-induced activity of the Cre-ER fusion protein (encoded by the *CreED-30* transgene) was verified by PCR genotyping⁴ before transplantation of the *Pax5*-deleted mature B cells.

FACS sorting of in vivo Pax5-deleted mature B cells. Lymph node cells from anti-IL-7Rα antibody-treated *Cd19-cre Pax5*^{fl/+} *Em-bcl2* *Ly5.2*⁺ mice were stained with biotinylated lineage marker antibodies (CD3ε, CD4, CD8α, CD11c, CD49b, Gr1, c-Kit, Mac1 and Ter119), and Lin⁺ non-B cells were depleted by magnetic cell sorting with streptavidin MACS beads (Miltenyi Biotec). *Pax5*-deficient mature B cells were sorted as Lin⁻ CD25⁺ IgM⁺ IgD⁻ cells⁴ after staining with cychrome-anti-CD25, PE-anti-IgM and APC-anti-IgD antibodies in combination with PE-Cy7-streptavidin staining of the Lin⁺ non-B cells. Flow cytometric reanalysis of the sorted *Pax5*-deficient B cells indicated a purity of more than 95% (Supplementary Fig. 6b).

Transplantations of mature Pax5-deleted B cells. *In vitro Pax5*-deleted mature B cells were washed in PBS and injected into the tail vein of 8–12-week-old *Rag2*^{-/-} *Ly5.1*⁺ or *Tcrα*^{-/-} recipient mice (2–4 × 10⁶ cells per mouse) that had been γ-irradiated with a sublethal dose (4 Gy) one day previously. *In vivo Pax5*-deleted mature B cells were similarly injected into *Rag2*^{-/-} *Ly5.1*⁺ mice (5 × 10⁵ cells per mouse). The drinking water for all reconstituted mice contained 1.14 g l⁻¹ neomycin.

Sorting of donor-derived pro-B and T cells. Dedifferentiated pro-B cells of donor origin were sorted from the bone marrow of transplanted *Rag2*^{-/-} *Ly5.1*⁺ mice as *Ly5.2*⁺ CD19⁺ B220⁺ c-Kit⁺ cells after staining with FITC-anti-*Ly5.2*, PE-anti-CD19, PE-Cy5-anti-B220 and APC-anti-cKit antibodies. Donor-derived double-positive thymocytes were sorted from the thymus of

transplanted *Rag2*^{-/-} *Ly5.1*⁺ mice as *Ly5.2*⁺ CD4⁺ CD8α⁺ cells after staining with FITC-anti-*Ly5.2*, PE-anti-CD4 and APC-anti-CD8α antibodies.

V(D)J recombination analysis. Sorted cells as well as tumour cells were digested with proteinase K, and DNA was isolated by phenol extraction and ethanol precipitation. V(D)J rearrangements of the *Igh*, *Igk*, *Igl* and *Tcrb* loci were amplified by PCR using published primers^{9,35–37} as well as newly designed Vβ gene primers, which are listed in Supplementary Table 4. PCR products were separated on agarose gels and visualized by ethidium bromide staining except for the data shown in Supplementary Fig. 10, which were obtained by Southern blotting using a germline PCR fragment (amplified with the Dβ2 and Jβ2 primers, Supplementary Table 4) as a DNA probe. To determine the DNA sequences of individual V–(D)J rearrangements, the PCR fragments were isolated from the agarose gel and cloned into the pGEM-Teasy vector (Promega); the DNA inserts of at least ten clones corresponding to the same PCR fragment were then sequenced.

RT-PCR analysis of rearranged immunoglobulin gene transcripts. RNA was prepared from *in vitro* grown *Pax5*^{Δ/-} tumour cell lines using the TRIzol reagent (GIBCO-BRL), and was then treated with DNase I (Promega) to remove contaminating DNA. Complementary DNA was synthesized with random hexamers and superscript II reverse transcriptase (Invitrogen). Rearranged immunoglobulin transcripts were amplified by PCR using the primers listed in Supplementary Table 5. The PCR products were separated on agarose gels and visualized by staining with ethidium bromide.

Microarray analysis. For microarray analysis, total RNA (3–5 µg) prepared from *Pax5*^{+/+} pro-B cells was used to generate a Cy3-labelled cDNA probe by oligo-dT₁₅-primed reverse transcription in the presence of Cy3-dUTP. Likewise, total RNA from *Pax5*^{-/-} pro-B cells, *Pax5*^{-/-} *Rag2*^{-/-} pro-B cells or *Pax5*^{Δ/-} lymph node tumours were used to generate Cy5-labelled cDNA probes by reverse transcription in the presence of Cy5-dUTP. The Cy3- and Cy5-labelled probes were hybridized to the mouse 'lymphochip'³⁸, which was scanned using a GenePix 4000 Scanner (Axon Instruments Inc.). The scanned data were evaluated with the GenePix Pro 5 chip software.

Intracellular cytokine and Foxp3 staining. Thymocytes were resuspended in RPMI 1680 medium (containing 10% FCS and 1 mM glutamine) at a concentration of 4 × 10⁶ cells ml⁻¹ and incubated for 6 h in the presence of PMA (20 ng ml⁻¹, Sigma), ionomycin (1 µM, Invitrogen) and brefeldin A (0.01%, BD Pharmingen). The cells were then washed, incubated with CD16/CD32 Fc-block solution (PD Pharmingen) and stained with FITC-anti-TCR-β, PE-Cy5-anti-CD4 and APC-anti-CD8α antibodies. After washing, the cells were fixed with 4% formaldehyde in PBS, permeabilized with the Perm/Wash solution (BD Biosciences) and stained with PE-anti-TNFα (MP6-XT22) or anti-IL-2 (JE56-5H4) antibodies before flow cytometric analysis. PE-labelled rat IgG1 (R3-34) and IgG2b (A95-1, BD Pharmingen) were used as negative isotype controls. The expression of Foxp3 was similarly analysed by intracellular staining of thymocytes using the PE-anti-Foxp3 (FJK-16s) antibody of the PE-anti-mouse/rat Foxp3 staining kit from eBiosciences.

T-cell proliferation assay. An anti-CD3ε antibody (2C11) was used at a concentration of 100 µg ml⁻¹ in PBS to coat 96-well plates (50 µl well⁻¹) for 90 min at 37 °C followed by two washes with PBS. Thymocytes, which were resuspended at 10⁶ cells ml⁻¹ in RPMI 1680 medium (containing 10% FCS and 1 mM glutamine), were seeded into the coated 96-well plates (2 × 10⁵ cells well⁻¹), and RPMI 1680 culture medium containing 5 µg ml⁻¹ of the anti-CD28 antibody 37.51 was added to the wells. Culture medium without antibody was added to control wells lacking any coating. The cells were incubated for 4–7 days, and their proliferation was measured by thymidine incorporation after ³H-thymidine addition for the last 24 h.

Immunization and germinal-centre analysis. Sheep red blood cells were washed in PBS and resuspended at 1–2 × 10⁹ cells ml⁻¹ followed by intraperitoneal injection of 100 µl into control mice or *Tcrα*^{-/-} mice, which had been transplanted with *in vitro Pax5*-deleted mature B cells eight-weeks before. Ten days after immunization, the spleens were isolated, embedded in OCT compound (Sakura) and snap-frozen on dry ice. Cryosections of the spleen were stained with a FITC-anti-IgD antibody (11-26c.2a, BD Pharmingen) and biotinylated PNA (B-1075, Vector Laboratories). FITC-anti-IgD was detected with an alkaline phosphatase-coupled anti-FITC antibody (Roche), which was visualized by incubation with Fast Blue (Sigma). Biotinylated PNA was detected with horseradish peroxidase-conjugated streptavidin (Zymed) followed by incubation with diaminobenzidine (DAB; Sigma).

In vitro culture of tumour and pro-B cells. Transformed cells isolated from *Cd19-cre Pax5*^{fl/-} lymph node tumours were cultured on γ-irradiated ST2 feeder cells in Iscove's modified Dulbecco's medium containing 2% heat-inactivated FCS, 0.03% (w/v) primatone RL (Quest International), 1 mM glutamine, 50 mM β-mercaptoethanol, 1% supernatant of IL-7-secreting J558L cells, 2.5% supernatant of Flt3L-producing SP2.0 cells and 2% supernatant of stem cell

factor-secreting CHO cells. Ly5.2⁺c-Kit⁺B220⁺CD19⁻ pro-B cells were sorted from the bone marrow of *Rag2*^{-/-} Ly5.1⁺ mice transplanted with *in vitro* Pax5-deleted mature B cells and were grown under the same conditions described above for the tumour cells.

Spectral karyotype analysis. Exponentially growing Pax5^{Δ/-} tumour cells were incubated for 5 h with 50 nM colcemid (Karyo-Max, GIBCO-BRL) in the absence of ST2 feeder cells. Metaphase chromosome spreads, prepared according to standard protocols³⁹, were hybridized using a SkyPaint DNA kit (Applied Spectral Imaging) according to the manufacturer's instructions. Spectral images were recorded and analysed using an interferometer and software from Applied Spectral Imaging⁴⁰.

Histological analysis. Organs of moribund *Cd19-cre Pax5*^{Δ/-} mice were fixed at 4 °C overnight with 4% formaldehyde in PBS followed by paraffin embedding, sectioning and staining with haematoxylin and eosin.

31. Urbánek, P., Wang, Z.-Q., Fetka, I., Wagner, E. F. & Busslinger, M. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell* **79**, 901–912 (1994).
32. Shinkai, Y. *et al.* RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* **68**, 855–867 (1992).
33. Rickert, R. C., Rajewsky, K. & Roes, J. Impairment of T-cell-dependent B-cell responses and B-1 cell development in CD19-deficient mice. *Nature* **376**, 352–355 (1995).
34. Strasser, A., Harris, A. W. & Cory, S. *bcl-2* transgene inhibits T cell death and perturbs thymic self-censorship. *Cell* **67**, 889–899 (1991).
35. Schlissel, M. S., Corcoran, L. M. & Baltimore, D. Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription. *J. Exp. Med.* **173**, 711–720 (1991).
36. Shaw, A. C., Swat, W., Davidson, L. & Alt, F. W. Induction of Ig light chain gene rearrangement in heavy chain-deficient B cells by activated Ras. *Proc. Natl Acad. Sci. USA* **96**, 2239–2243 (1999).
37. Wolfer, A., Wilson, A., Nemir, M., MacDonald, H. R. & Radtke, F. Inactivation of *Notch1* impairs VDJ β rearrangement and allows pre-TCR-independent survival of early $\alpha\beta$ lineage thymocytes. *Immunity* **16**, 869–879 (2002).
38. Shaffer, A. L. *et al.* XBPI, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* **21**, 81–93 (2004).
39. Barch, M. J., Knutsen, T. & Spurbeck, J. L. *The AGT Cytogenetic Laboratory Manual* (Lippincott, Philadelphia, 1997).
40. Franco, S. *et al.* H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. *Mol. Cell* **21**, 201–214 (2006).

LETTERS

IgH class switching and translocations use a robust non-classical end-joining pathway

Catherine T. Yan^{1,2,3,4}, Cristian Boboila^{1,2,3,4}, Ellen Kris Souza^{1,2,3,4}, Sonia Franco^{1,2,3,4}, Thomas R. Hickernell^{1,2,3,4}, Michael Murphy^{1,2,3,4}, Sunil Gumaste^{1,2,3,4}, Mark Geyer², Ali A. Zarrin^{1,2,3,4}, John P. Manis^{2,5}, Klaus Rajewsky^{3,5} & Frederick W. Alt^{1,2,3,4}

Immunoglobulin variable region exons are assembled in developing B cells by V(D)J recombination. Once mature, these cells undergo class-switch recombination (CSR) when activated by antigen. CSR changes the heavy chain constant region exons (CH) expressed with a given variable region exon from C μ to a downstream CH (for example, C γ , C ϵ or C α), thereby switching expression from IgM to IgG, IgE or IgA. Both V(D)J recombination and CSR involve the introduction of DNA double-strand breaks and their repair by means of end joining^{1,2}. For CSR, double-strand breaks are introduced into switch regions that flank C μ and a downstream CH, followed by fusion of the broken switch regions¹. In mammalian cells, the 'classical' non-homologous end joining (C-NHEJ) pathway repairs both general DNA double-strand breaks and programmed double-strand breaks generated by V(D)J recombination^{2,3}. C-NHEJ, as observed during V(D)J recombination, joins ends that lack homology to form 'direct' joins, and also joins ends with several base-pair homologies to form microhomology joins^{3,4}. CSR joins also display direct and microhomology joins, and CSR has been suggested to use C-NHEJ^{5–8}. Xrcc4 and DNA ligase IV (Lig4), which cooperatively catalyse the ligation step of C-NHEJ, are the most specific C-NHEJ factors; they are absolutely required for V(D)J recombination and have no known functions other than C-NHEJ². Here we assess whether C-NHEJ is also critical for CSR by assaying CSR in Xrcc4- or Lig4-deficient mouse B cells. C-NHEJ indeed catalyses CSR joins, because C-NHEJ-deficient B cells had decreased CSR and substantial levels of IgH locus (immunoglobulin heavy chain, encoded by *Igh*) chromosomal breaks. However, an alternative end-joining pathway, which is markedly biased towards microhomology joins, supports CSR at unexpectedly robust levels in C-NHEJ-deficient B cells. In the absence of C-NHEJ, this alternative end-joining pathway also frequently joins *Igh* locus breaks to other chromosomes to generate translocations.

We generated Xrcc4 (X-ray repair complementing defective repair in Chinese hamster cells 4)-deficient B cells by two different and complementary methods to circumvent the fact that Xrcc4 is absolutely required for generation of mature B cells owing to its requirement for V(D)J recombination⁹. One method involved generating Xrcc4-deficient mice (Xrcc4^{-/-})^{9,10} that harboured pre-assembled ('knocked in') immunoglobulin heavy chain (IgH, B1-8-HC) and immunoglobulin light chain (IgL, 3–83k-LC) variable-region exons¹¹ and that were heterozygous or homozygous for a p53 null mutation to rescue neuronal apoptosis and embryonic lethality of the Xrcc4^{-/-} genotype¹². Because the p53^{+/-} and p53^{-/-} backgrounds gave similar results, we collectively refer to these mice as XP-T/HL (where T indicates targeted germline knockout, H indicates heavy chain

knockin, and L indicates light chain knockin). We used the germline gene-inactivation approach to ensure complete absence of Xrcc4 in mature B cells. XP-T/HL mice contained populations of surface IgM⁺ splenic B cells (Fig. 1a); however, as expected^{5,11}, B-cell numbers were relatively low (see Methods) and there were no T cells (data not shown). Therefore, we also inactivated Xrcc4 specifically in mature B cells by generating mice that harboured one copy of a loxP-flanked (floxed) Xrcc4 allele (Xrcc4^f)¹⁰ and one copy of an Xrcc4 null allele (Xrcc4⁻) plus a transgene that drives Cre recombinase expression in later stages of the B lineage from a CD21 promoter¹³ (termed CD21-cre-Xrcc4^{f/-} mice). CD21-cre-Xrcc4^{f/-} mice had normal numbers of IgM⁺ splenic B and T cells (not shown). We were unable to detect intact Xrcc4 genes, transcripts or proteins in cultured CD21-cre-Xrcc4^{f/-} B cells (Supplementary Figs 1–4).

To test the role of Xrcc4 in proliferation and CSR in activated B cells, XP-T/HL, CD21-cre-Xrcc4^{f/-} or control (HL, Xrcc4^{+/+}-HL or Xrcc4^{f/f}) splenic B220⁺CD43⁻ B cells were cultured for up to 5 days with either anti-CD40 plus interleukin (IL)-4 to stimulate CSR to IgG1 and IgE or bacterial lipopolysaccharide (LPS) to stimulate CSR to IgG2b and IgG3. Proliferation of XP-T/HL and CD21-cre-Xrcc4^{f/-} B cells was not markedly impaired (Fig. 1b), even though both mutants were highly sensitive to ionizing radiation (Fig. 1c), the latter being consistent with their complete deficiency of Xrcc4. Enzyme-linked immunosorbent assays (ELISAs) for secretion of IgH isotypes in appropriately stimulated XP-T/HL and CD21-cre-Xrcc4^{f/-} B cells revealed substantial production of the IgH isotypes assayed (IgG1, IgG2b, IgG3 and IgE) (Supplementary Fig. 5 and Supplementary Table 1). In addition, flow cytometry assays for surface IgG1, IgG2b and IgG3 also revealed substantial switching by appropriately stimulated XP-T/HL or CD21-cre-Xrcc4^{f/-} B cells (Fig. 2a and Supplementary Table 2). On the basis of ten different experiments, average levels of IgG1 were at about 25% of those of controls, but sometimes were as high as 50% (Supplementary Table 2). A time course of IgG1 switching in CD21-cre-Xrcc4^{f/-} B cells revealed comparable relative IgH switching between mutant and control B cells at 2.5 days when switching was well below peak levels in controls, and at late time points (for example, 4.5 days) when IgH switching was maximal (Supplementary Fig. 6). An enzyme-linked immunospot (ELISPOT) assay, which provides an assessment of the frequency of antibody-secreting cells at the single-cell level¹⁴, revealed switching to IgG1 to occur at 20–50% of control levels and switching to IgG3 at nearly 50% of control levels for XP-T/HL and CD21-cre-Xrcc4^{f/-} B cells, as opposed to the negligible switching in cells deficient for activation-induced cytidine deaminase (AID) (Fig. 2b and Supplementary Table 3). Therefore, all three assays demonstrated that Xrcc4-deficient B cells have decreased overall

¹Howard Hughes Medical Institute, ²The Children's Hospital, ³Immune Disease Institute, ⁴Department of Genetics, ⁵Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA.

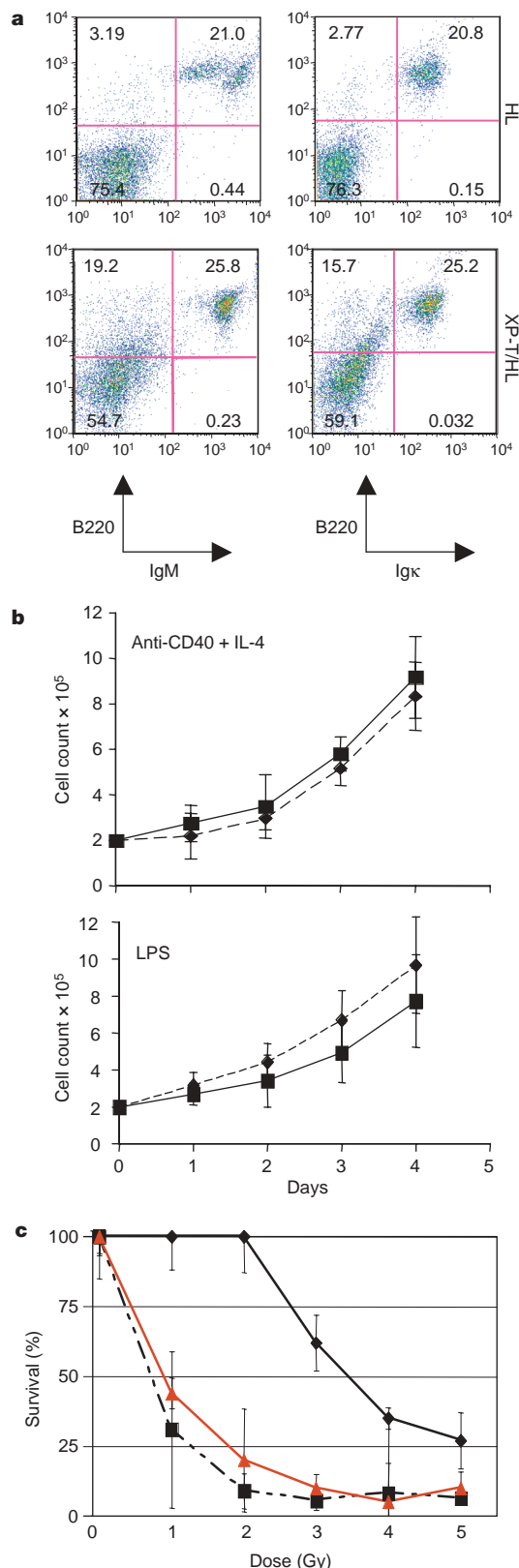


Figure 1 | Proliferation and ionizing radiation sensitivity of *Xrcc4*-deficient B cells. **a**, Surface IgM⁺ and Igk⁺ expression on XP-T/HL and HL splenocytes assayed by flow cytometry. **b**, Daily cell counts of activated XP-T/HL (solid line) and HL (dashed line) B cells for three independent experiments. *Xrcc4*^{cl} and *CD21-cre-Xrcc4*^{cl} B-cell cultures gave similar results (not shown). Data are presented as mean ± s.e.m. **c**, Ionizing-radiation sensitivity of *Xrcc4*^{cl} (diamonds), *CD21-cre-Xrcc4*^{cl} (squares) and XP-T/HL (triangles) B cells activated with anti-CD40 plus IL-4. Representative results of multiple experiments are shown as average per cent survival ± s.e.m. Ionizing-radiation sensitivity of HL cells (not shown) was equivalent to results for *Xrcc4*^{cl} B cells.

levels of IgH class switching but still carry out this chromosomal end-joining process at robust levels.

NHEJ is catalysed specifically by Lig4, with Xrcc4 serving as a requisite cofactor for activity and stability of the Lig4 protein¹⁵. Indeed, assays for Lig4 protein in *Xrcc4*-deficient cells revealed greatly reduced levels (Supplementary Fig. 7), consistent with the identical phenotypes of *Xrcc4*- and Lig4-deficient cells and mice^{9,15,16}. To confirm unequivocally that the high level of CSR observed in *Xrcc4*-deficient B cells was not caused by some unanticipated Lig4 (and NHEJ) activity, we bred the homozygous Lig4-inactivating mutation¹⁶ into the P-T/HL background to generate LP-T/HL B cells (Supplementary Fig. 8a). The Lig4 mutation deletes much of the *Lig4* gene including the portion encoding the active site of the protein¹⁶. Notably, LP-T/HL B cells activated with anti-CD40 plus IL-4 still switched to IgG1, as assayed by both surface staining (Fig. 2c and Supplementary Fig. 8b) and ELISPOT (Supplementary Fig. 9), at levels up to 50% of controls. Likewise, the LP-T/HL B cells also underwent switching to IgE (Supplementary Fig. 8c). Therefore, we conclude that there is a robust *Xrcc4*- and Lig4-independent CSR pathway. In this regard, B cells deficient for Ku70 and Ku80—C-NHEJ factors also required for V(D)J recombination²—have severe proliferation defects and a complete inability to undergo CSR^{5–7}. Our current findings with *Xrcc4*- and Lig4-deficient B cells suggest that the more severe phenotypes of the Ku70- and Ku80-deficient cells might reflect functions of Ku proteins in processes other than C-NHEJ¹.

We used fluorescence *in situ* hybridization (FISH) to test whether *Xrcc4* deficiency resulted in general and *Igh*-locus-specific chromosomal breaks in XP-T/HL and *CD21-cre-Xrcc4*^{cl} B cells, as compared to control B cells stimulated with anti-CD40 plus IL-4. On the basis of a telomere FISH (T-FISH) assay¹⁷, both XP-T/HL and *CD21-cre-Xrcc4*^{cl} B cells exhibited increased levels of general chromosomal aberrations, predominantly chromosome breaks (Fig. 3 and Supplementary Tables 4 and 5); this is consistent with a defect in the repair of double-strand breaks (DSBs) mostly initiated in a pre-replicative cell-cycle phase. To assay for breaks or translocations within the *Igh* locus, we used a FISH assay with one probe that recognizes sequences just upstream of the *Igh* V_H domain on the telomeric region of chromosome 12 (5' *Igh* probe) plus a probe that recognizes sequences immediately downstream of the CH locus (3' *Igh* probe)¹⁷ (Fig. 3a). This assay revealed chromosomal breaks within the *Igh* locus, visualized as separated 3' and 5' *Igh* signals, in a large proportion (20–30% on average) of metaphases from activated XP-T/HL and *CD21-cre-Xrcc4*^{cl} B cells, in contrast to low levels in the control cells (Fig. 3 and Supplementary Tables 4 and 5). Additional FISH analyses mapped many of the *Igh* locus breaks in *CD21-cre-Xrcc4*^{cl} B cells to a region downstream of C_μ and upstream of C_α, consistent with an origin during attempted CSR (Supplementary Fig. 10). These findings demonstrate that *Xrcc4* is required for repair of a subset of *Igh* locus breaks that occur during CSR. Notably, a significant proportion (over 20% on average) of *Igh* locus breaks that did not participate in CSR in XP-T/HL and *CD21-cre-Xrcc4*^{cl} B cells participated in translocations (Fig. 3d, e and Supplementary Table 5). Thus, in the absence of *Xrcc4*, another pathway in B cells activated for CSR robustly catalyses translocations of unrepaired *Igh* locus breaks.

End-joining assays on transfected DNA substrates implicated an alternative end-joining pathway that is biased towards microhomology and operates independently of known C-NHEJ factors^{4,18–21}. Alternative end-joining also has been implicated in oncogenic translocations in *Xrcc4*- or Lig4-deficient cells²². However, alternative end-joining does not rescue V(D)J recombination in the absence of C-NHEJ factors^{9,16} and it has been thought that it does not function efficiently on chromosomal DSBs^{4,18,19}. Human B cells that harboured hypomorphic LIG4 mutations showed considerably increased use of microhomology at S_μ to S_α junctions, but not at S_μ to S_γ junctions; this might reflect either increased use of an alternative end-joining pathway or altered activity of residual LIG4

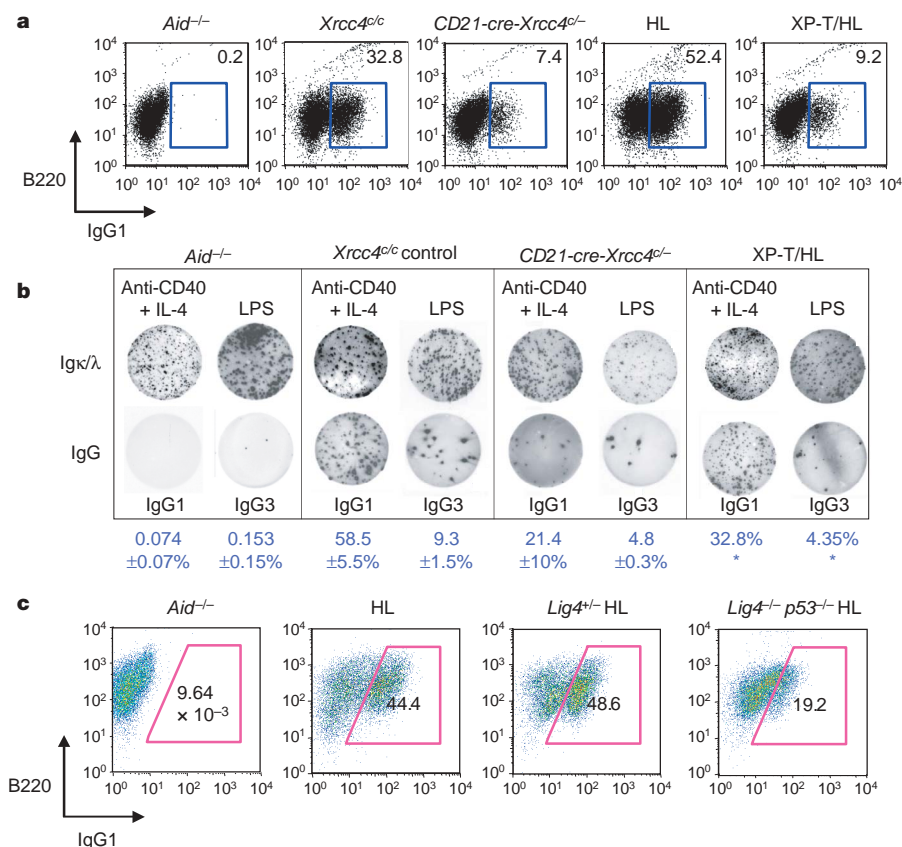


Figure 2 | CSR in Xrcc4-deficient B cells.

a, Surface IgG1 expression of indicated CD43⁺ splenic B cells on day four of anti-CD40 plus IL-4 stimulation. (Details of all experiments are in Supplementary Table 2; experiment 7 is shown.) **b**, ELISPOT analysis of switching to IgG1 and IgG3 in Xrcc4-deficient B cells. Average per cent ± s.d. are shown below each panel. Three mice were used for each experiment, except where denoted with an asterisk (in XP-T/HL) where two mice were used. (See Supplementary Table 3 for details.) **c**, Flow cytometric analysis of surface IgG1 expression on day-four anti-CD40- and IL-4-activated Lig4-deficient and control CD43⁺ splenic B cells (details in Supplementary Fig. 8). AID, activation-induced cytidine deaminase (also known as Aicda).

and C-NHEJ⁸. To determine the nature of the DSB repair pathway that catalyses CSR and translocations in activated XP-T/HL and CD21-cre-Xrcc4^{cl/c} B cells, we analysed CSR junctions. We activated XP-T/HL and control cells with either anti-CD40 plus IL-4 or LPS and then used polymerase chain reaction (PCR) to isolate, respectively, Sμ to Sγ1 and Sμ to Sε junctions or Sμ to Sγ2b and Sμ to Sγ3 junctions (Fig. 4, Supplementary Table 6 and Supplementary Figs

11–14). We also isolated Sμ to Sγ1 junctions from CD21-cre-Xrcc4^{cl/c} B cells and hybridomas stimulated by anti-CD40 plus IL-4 (Supplementary Fig. 15 and Supplementary Table 6). Comparison of CSR junctions from Xrcc4-deficient B cells to those of controls revealed a profound difference. HL control (or Xrcc4^{+/+} HL) sequences contained about 30–60% direct joins, depending on the downstream switch region, with most of the remainder being short

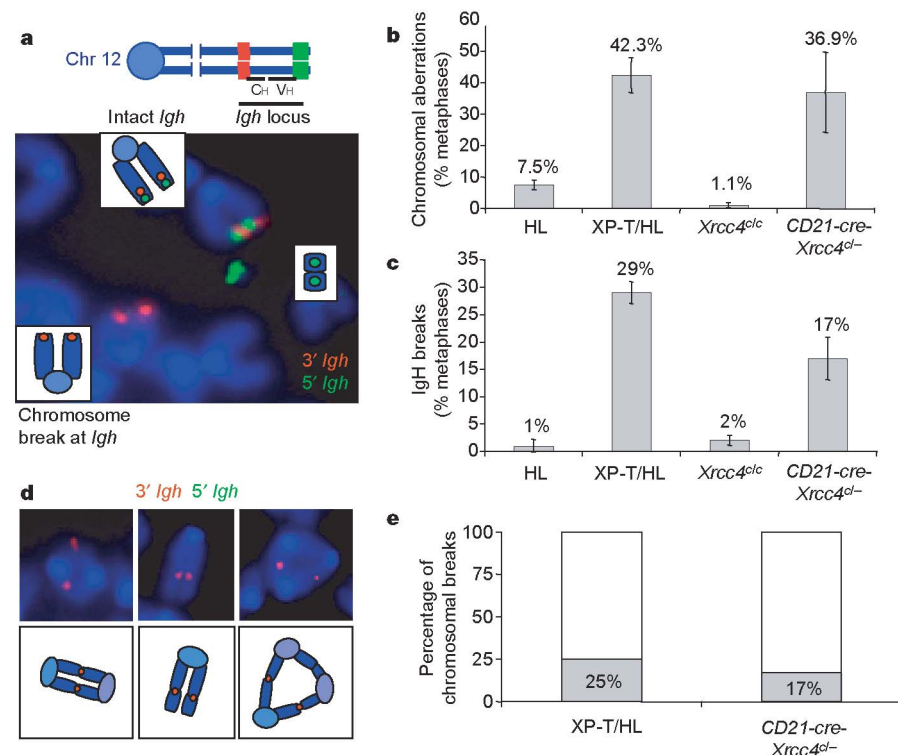


Figure 3 | Frequent general and *Igh*-locus-specific chromosomal breaks and translocations in Xrcc4-deficient B cells. Metaphases from anti-CD40- and IL-4-activated B cells (day four) were analysed by IgH FISH (**a**, **c–e**) or T-FISH (**b**) (details in Supplementary Tables 4 and 5). **a**, Representative *Igh* locus breaks in XP-T/HL cultures. Intact *Igh* locus on chromosome 12 shows red and green signals; broken loci show only red or split red and green. **b**, Quantification of general chromosomal breaks. **c**, Quantification of *Igh* locus breaks. Data in **b** and **c** are presented as mean ± s.e.m. **d**, Representative *Igh* locus translocations. **e**, Per cent untranslocated free (clear bars) versus translocated (filled bars) *Igh* locus chromosomal breaks.

(1–4 bp) microhomology joins (Fig. 4 and Supplementary Table 6). In contrast, no direct CSR joins were found in more than 100 junctions from *Xrcc4*-deficient B cells, regardless of the downstream switch region used (Fig. 4, Supplementary Table 6 and Supplementary Figs 11–15). Most *Xrcc4*-deficient CSR joins had microhomology ($P = < 0.0001$) and the remainder contained insertions of various lengths (Supplementary Table 6) and, as such, also might have been microhomology-mediated. Microhomology joins from *Xrcc4*-deficient B cells showed an increased distribution of junctions with longer microhomologies than those of the wild type (Fig. 4a–d and Supplementary Table 6), but the distribution of breakpoints within the switch regions and levels of point mutations were similar to those of controls (Supplementary Fig. 16 and Supplementary Table 7). In preliminary studies, a set of analysed *Lig4*-deficient CSR junctions showed a similar lack of direct joins and a bias towards microhomology joins (C.B., C.T.Y. and F.A., unpublished observations).

We show that an alternative end-joining pathway can join chromosomal DSBs in the physiological context of CSR. Unlike C-NHEJ, this alternative end-joining pathway operates in the absence of *Xrcc4* and *Lig4* and does not generate readily detectable levels of direct CSR joins. Thus, alternative end-joining is mechanistically distinct from C-NHEJ, both in the use of different ends and

of a different ligase, with *Lig3* being a likely candidate^{18,23}. Until now it has been concluded that alternative end-joining does not function efficiently on chromosomal DSBs, even in the absence of C-NHEJ^{4,18,19}. However, we find that this pathway is much more robust in chromosomal DSB repair than previously expected, at least during CSR. Thus, in the absence of C-NHEJ, alternative end-joining catalyses substantial CSR, even early after B-cell activation when CSR is still on the rise. Yet, *Xrcc4*-deficient B cells are sensitive to ionizing radiation and have increased general genomic instability; this is consistent with our finding that alternative end-joining does not repair all CSR-associated DSBs. It is possible that particular ends favour alternative end-joining, such as those with longer microhomology. Notably, some of the $S\mu$ to $S\alpha$ junctions in normal human B cells⁸ and $S\mu$ to $S\epsilon$ junctions in normal mouse B cells (Supplementary Fig. 14) use long microhomology not characteristic of C-NHEJ^{24,25}. Thus, alternative end-joining may compete better with C-NHEJ for CSR events involving switch regions that are relatively homologous in repeat structure (for example, $S\mu$, $S\epsilon$ and $S\alpha$) and which, thereby, provide more optimal microhomology substrates⁸. Furthermore, the increased frequency of CSR junctions with longer microhomologies in certain mismatch-repair-deficient mice^{26,27} might reflect alterations in optimal substrate availability that favour alternative end-joining. Our findings raise the question of why alternative

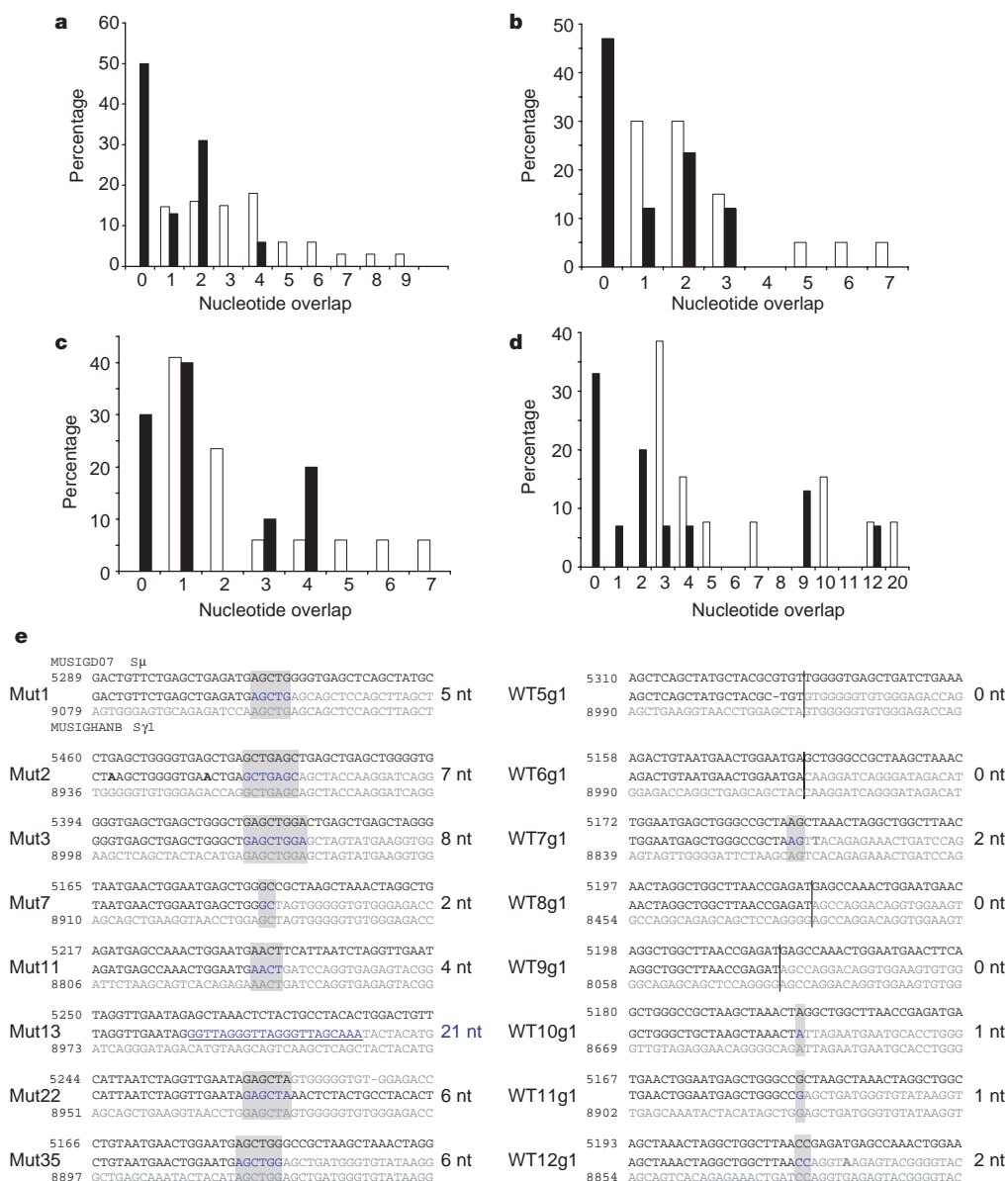


Figure 4 | Structure of *Xrcc4*-deficient CSR junctions.

a–d, Percentage of sequences with indicated nucleotide (nt) overlap for $S\mu$ and $S\gamma 1$ (**a**), $S\mu$ and $S\gamma 2b$ (**b**), $S\mu$ and $S\gamma 3$ (**c**) and $S\mu$ and $S\epsilon$ (**d**) junctions (details in Supplementary Table 6 and Supplementary Figs 11–15). XP-T/HL, open bars; HL, filled bars. **e**, HL (wild type) and XP-T/HL $S\mu$ – $S\gamma 1$ junctions (additional sequences in Supplementary Fig. 11). Microhomology is identified as the longest region with perfect homology. GenBank annotated sequence²⁷ MUSIGD07 ($S\mu$, top row); CSR junctions (middle row); GenBank sequence MUSIGHANB ($S\gamma 1$, bottom row). Vertical line, no sequence identity; blue text with grey highlight, microhomology; underlined blue text, insertions; bold text, point mutations; and dash, deletions. Nucleotide overlaps (black) and insertions (blue) are shown on the right of each sequence.

end-joining does not rescue any V(D)J joining in developing *Xrcc4*- or *Lig4*-deficient lymphocytes, despite 50% of normal V(D)J joins having short microhomologies²⁸. The answer seems to be provided by recent findings that RAG proteins channel V(D)J DSBs into C-NHEJ and exclude alternative end-joining³¹. In contrast, our current findings show that AID-initiated DSBs during CSR are not similarly restricted to C-NHEJ. Finally, in the absence of C-NHEJ, we find that *Igh* locus CSR breaks are very frequently joined to other chromosomes to generate translocations, which also are probably catalysed largely by alternative end-joining²². In this context, alternative end-joining may be inherently 'error-prone'^{23,29,30}, with this feature being enhanced by frequent occurrence of unrepaired DSBs in the C-NHEJ-deficient background.

METHODS SUMMARY

Generation of XP-T/HL, *CD21-cre-Xrcc4*^{fl} and LP-T/HL mice. XP-T/HL mice (*Xrcc4*^{-/-}*p53*^{+/-}HC/LC and *Xrcc4*^{-/-}*p53*^{-/-} HC/LC) on the 129 Sv/ev background were generated by breeding *Xrcc4*^{+/-}*p53*^{+/-} with HC/LC (HL) knock-in mice¹¹. HL and *Xrcc4*^{+/-} HL mice were generated as control cohorts. *CD21-cre-Xrcc4*^{fl} and *CD21-cre-Xrcc4*^{fl}*p53*^{+/-} mice were generated by crossing *CD21-cre* transgenic mice¹³ harbouring heterozygous deletions of *Xrcc4* and *p53* (*CD21-cre-Xrcc4*^{+/-}*p53*^{+/-}) with *Xrcc4* floxed (*Xrcc4*^{fl}) mice¹⁰. LP-T/HL mice (*Lig4*^{-/-}*p53*^{+/-}HC/LC) and *Lig4*^{+/-}HL mice were generated by crossing *Lig4*^{+/-}*p53*^{+/-} with HL mice. Because of the limited lifespan of the *Xrcc4*^{-/-}*p53*^{+/-} genetic background, all XP-T/HL and LP-T/HL experiments were performed at 3–5 weeks of age. *CD21-cre-Xrcc4*^{fl} and *Xrcc4*^{fl} mice were analysed at 6–10 weeks of age.

Mature B-cell purification, *in vitro* culture conditions and CSR assays. CD43⁺ B cells were isolated and cultured as previously described¹⁷. Cells were maintained daily at 5×10^5 cells ml⁻¹ and were sampled on various days for RNA, DNA, protein, flow cytometry¹⁷, ELISA¹¹ and ELISPOT¹⁴ analysis. RNA isolation, Southern blotting and western blotting were performed as previously described¹⁰.

CSR junction analyses. μ - γ 1, μ - γ 2b and μ - γ 3 junctions were amplified by PCR from genomic DNA prepared from splenic B cells activated with anti-CD40 plus IL-4 for 4 days, as described²⁷. At least three independent reactions each from three control and three experimental mice were used to obtain products that hybridized with the γ 1 probe; two mice each were used for other downstream switch regions. PCR products were cloned into the Topo-TA (Invitrogen) or the pGEM-T vector (Promega), sequenced and analysed using Lasergene software and the NCBI database. The GenBank sequence files used were as follows: MUSIGD07 (for μ), MUSIGHANB (for γ 1), MUSIGD18 (for γ 3), MMU85373 (for γ 2b) and MUSIGHAHX (for γ 2). γ 2 primers for μ - γ 2 junctions were: Se1, CAAGTTAAGCTCATGCCAGGTCAA; and Se2, CCAGGCCAGGCCAGTCTGCTCAGT. PCR was performed using μ - γ conditions²⁷. **Two-colour FISH and telomere-FISH.** Metaphase spreads from day-four anti-CD40 plus IL-4 B-cell cultures were prepared. Two-colour FISH to detect *Igh*-locus-specific chromosomal aberrations and telomere staining (T-FISH) to detect general aberrations were performed as previously described¹⁷.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 April; accepted 11 June 2007.

Published online 22 August 2007.

1. Chaudhuri, J. *et al.* Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.* **94**, 157–214 (2007).
2. Rooney, S., Chaudhuri, J. & Alt, F. W. The role of the non-homologous end-joining pathway in lymphocyte development. *Immunol. Rev.* **200**, 115–131 (2004).
3. Roth, D. B. Restraining the V(D)J recombinase. *Nature Rev. Immunol.* **3**, 656–666 (2003).
4. Verkaik, N. S. *et al.* Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. *Eur. J. Immunol.* **32**, 701–709 (2002).
5. Manis, J. P. *et al.* Ku70 is required for late B cell development and immunoglobulin heavy chain class switching. *J. Exp. Med.* **187**, 2081–2089 (1998).
6. Casellas, R. *et al.* Ku80 is required for immunoglobulin isotype switching. *EMBO J.* **17**, 2404–2411 (1998).
7. Reina-San-Martin, B. *et al.* H2AX is required for recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation. *J. Exp. Med.* **197**, 1767–1778 (2003).
8. Pan-Hammarstrom, Q. *et al.* Impact of DNA ligase IV on nonhomologous end joining pathways during class switch recombination in human cells. *J. Exp. Med.* **201**, 189–194 (2005).

9. Gao, Y. *et al.* A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* **95**, 891–902 (1998).
10. Yan, C. T. *et al.* XRCC4 suppresses medulloblastomas with recurrent translocations in p53-deficient mice. *Proc. Natl Acad. Sci. USA* **103**, 7378–7383 (2006).
11. Manis, J. P., Dudley, D., Kaylor, L. & Alt, F. W. IgH class switch recombination to IgG1 in DNA-PKcs-deficient B cells. *Immunity* **16**, 607–617 (2002).
12. Gao, Y. *et al.* Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. *Nature* **404**, 897–900 (2000).
13. Kraus, M., Alimzhanov, M. B., Rajewsky, N. & Rajewsky, K. Survival of resting mature B lymphocytes depends on BCR signaling via the Ig α / β heterodimer. *Cell* **117**, 787–800 (2004).
14. Zarrin, A. A. *et al.* Antibody class switching mediated by yeast endonuclease-generated DNA breaks. *Science* **315**, 377–381 (2007).
15. Bryans, M., Valenzano, M. C. & Stamato, T. D. Absence of DNA ligase IV protein in XR-1 cells: evidence for stabilization by XRCC4. *Mutat. Res.* **433**, 53–58 (1999).
16. Frank, K. M. *et al.* Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* **396**, 173–177 (1998).
17. Franco, S. *et al.* H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. *Mol. Cell* **21**, 201–214 (2006).
18. Wang, H. *et al.* DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res.* **65**, 4020–4030 (2005).
19. Roth, D. B. Amplifying mechanisms of lymphomagenesis. *Mol. Cell* **10**, 1–2 (2002).
20. Kabotianski, E. B., Gomelsky, L., Han, J. O., Stamato, T. D. & Roth, D. B. Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res.* **26**, 5333–5342 (1998).
21. Guirouilh-Barbat, J. *et al.* Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Mol. Cell* **14**, 611–623 (2004).
22. Zhu, C. *et al.* Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell* **109**, 811–821 (2002).
23. Audebert, M., Salles, B. & Calsou, P. Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem.* **279**, 55117–55126 (2004).
24. Daley, J. M. & Wilson, T. E. Rejoining of DNA double-strand breaks as a function of overhang length. *Mol. Cell. Biol.* **25**, 896–906 (2005).
25. Gu, J. *et al.* XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *EMBO J.* **26**, 1010–1023 (2007).
26. Schrader, C. E., Vardo, J. & Stavnez, J. Role for mismatch repair proteins Msh2, Mlh1, and Pms2 in immunoglobulin class switching shown by sequence analysis of recombination junctions. *J. Exp. Med.* **195**, 367–373 (2002).
27. Ehrenstein, M. R., Rada, C., Jones, A. M., Milstein, C. & Neuberger, M. S. Switch junction sequences in PMS2-deficient mice reveal a microhomology-mediated mechanism of Ig class switch recombination. *Proc. Natl Acad. Sci. USA* **98**, 14553–14558 (2001).
28. Komori, T., Okada, A., Stewart, V. & Alt, F. W. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. *Science* **261**, 1171–1175 (1993).
29. Chen, S. *et al.* Accurate *in vitro* end joining of a DNA double strand break with partially cohesive 3'-overhangs and 3'-phosphoglycolate termini: effect of Ku on repair fidelity. *J. Biol. Chem.* **276**, 24323–24330 (2001).
30. Weinstock, D. M., Richardson, C. A., Elliott, B. & Jasim, M. Modeling oncogenic translocations: distinct roles for double-strand break repair pathways in translocation formation in mammalian cells. *DNA Repair (Amst.)* **5**, 1065–1074 (2006).
31. Corneo, B. *et al.* Rag mutations reveal robust alternative end joining. *Nature* doi:10.1038/nature06168 (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Schatz for the gift of the Lig4 antibody, M. Alimzhanov, S. Casola and J.-B. Telliez for technical assistance and suggestions, and members of the Alt laboratory for discussions. This work was supported by NIH grants (to F.W.A.), an NIH postdoctoral training grant (to C.T.Y.) and a Long-Term Fellowship of the European Molecular Biology Organization (to S.F.). F.W.A. is an investigator of the Howard Hughes Medical Institute.

Author Contributions F.W.A. and C.T.Y. planned the studies and analysed and interpreted the data. C.T.Y. generated the reagents and performed or oversaw all the experiments described. E.K.S. along with C.T.Y. purified and stimulated XP-T/HL B cells for the studies described. C.B. with C.T.Y. generated and analysed the LP-T/HL mice. E.K.S. generated and analysed (with C.T.Y.) most of the switch junctions and performed all statistical analysis. C.T.Y., T.R.H., S.F. and S.G. analysed metaphases for telomere-FISH and IgH FISH. C.T.Y., T.R.H., E.K.S. and C.B. performed ELISPOT analysis with technical assistance and reagents from A.A.Z. C.B. with E.K.S. isolated and analysed hybridoma sequence junctions; J.P.M. with M.G. provided expertise in B-cell stimulation, fluorescence-activated cell sorting and analysis of sequence junctions; and K.R. provided *CD21-cre* and HL mice and helped interpret data. F.W.A. and C.T.Y. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.W.A. (alt@enders.tch.harvard.edu).

METHODS

Generation of XP-T/HL, *CD21-cre-Xrcc4*^{Δ/Δ} and LP-T/HL mice. All mice are in the 129 Sv/ev genetic background. The various genotypes were identified by Southern blotting as described¹⁰. Live B220⁺/IgM⁺/Igκ⁺ peripheral B-cell counts in mice were as follows: $1.2 \pm 0.2 \times 10^6$ in *Xrcc4*^{Δ/Δ}*p53*^{+/Δ} HL mice ($n = 8$), $31 \pm 4 \times 10^6$ in HL mice ($n = 7$), $29 \pm 4 \times 10^6$ in *Xrcc4*^{+/Δ} HL mice ($n = 2$), and 20×10^6 in *Xrcc4*^{Δ/Δ}*p53*^{Δ/Δ} HL mice ($n = 1$). Live B220⁺/IgM⁺/Igκ⁺ B-cell counts for the two LP-T/HL mice analysed were $\sim 1.0 \pm 0.2 \times 10^6$.

Splenic B-cell cultures. Single-cell suspensions of fractionated CD43⁺ populations were cultured at 5×10^5 cells ml⁻¹ in RPMI medium supplemented with 10% FCS and 500 ng ml⁻¹ anti-CD40 antibody (Pharmingen) with 25 ng ml⁻¹ mouse recombinant IL-4 (R&D Systems) (or 25 ng ml⁻¹ LPS plus 25 ng ml⁻¹ IL-4) to assay for switching to IgG1 and IgE, or 25 ng ml⁻¹ LPS (Sigma) to assay for switching to IgG3 and IgG2b, as previously described¹⁷. Cultured cells were maintained daily at a density of 5×10^5 cells ml⁻¹, and sampled on various days for RNA, flow cytometry analyses, ELISA, protein and genomic DNA. Western blotting using a rabbit polyclonal affinity-purified antibody against murine Lig4 antibody³² was performed at 1:250 in 3% milk/TBS-T.

Cell proliferation and analysis of ionizing radiation sensitivity. Proliferation of *Xrcc4*-deficient B cells was analysed by daily cell counts of activated XP-T/HL versus HL mice; *Xrcc4*^{Δ/Δ} and *CD21-cre-Xrcc4*^{Δ/Δ} B220⁺ CD43⁺ splenic B cells were quantified by trypan blue to exclude dead cells. To determine whether *Xrcc4*-deficient B cells are sensitive to ionizing radiation, purified CD43⁺ B cells cultured in anti-CD40 plus IL-4 from a total of $n = 5$ *Xrcc4*^{Δ/Δ}, $n = 5$ *CD21-cre-Xrcc4*^{Δ/Δ}, $n = 2$ HL, $n = 1$ *Xrcc4*^{Δ/Δ}*p53*^{Δ/Δ} HL and $n = 2$ *Xrcc4*^{Δ/Δ}*p53*^{+/Δ} HL mice—one or more independent stimulations were performed for each mouse on day 2.5 of activation—were plated as triplicate samples, irradiated with a dose of 0–5 Gy, and counted with trypan blue to exclude dead cells 24 h after irradiation.

Flow cytometry analyses. For fluorescence-activated cell sorting analyses, cells were washed twice in PBS/2% FCS and stained with various antibodies conjugated with fluorescein (IgG1 and IgG2b), phycoerythrin (IgE, IgG3 and IgM) or Cy7-phycoerythrin (B220) as previously described¹⁷. The cells were analysed by using a FACSCalibur (BD Biosciences), and data were interpreted using FLOWJO software (Tree Star). At least 10,000 events of gated live, lymphoid cells, as determined by forward scatter versus side scatter, were recorded.

ELISA and ELISPOT. ELISAs to detect IgM and various IgG isotypes in cultured supernatants, obtained after 5 days of stimulation, were performed as described¹¹

using isotype-specific antibodies purchased from Southern Biotechnology Associates. Antibodies for IgE detection were purchased from BD Pharmingen. ELISPOTs were performed essentially as described¹⁴.

Two-colour FISH. Lymphocytes were incubated in colcemid (KaryoMAX, Gibco), swollen in pre-warmed 30 mM sodium citrate for 25 min at 37 °C (or 70 mM KCl for 15 min at 25 °C), fixed in methanol/acetic acid (3/1) and dropped onto slides that were then air-dried overnight. We detected metaphase chromosomes containing the 3' end of the *Igh* locus using BAC199 (a gift of B. Birshstein) and those containing the 5' end using BAC207 (a gift of P. Brodeur), and identified chromosomes containing the Cμ–Cδ portion of the locus using an isolated internal 10 kb Cμ/Cδ genomic probe. BACs were labelled with either biotin (biotin-nick translation mix, Roche) or digoxigenin (dig-nick translation mix, Roche), as per the manufacturer's instructions. Two-hundred nanograms of BAC DNA was precipitated with mouse *Cot1* DNA (Invitrogen; ratio of BAC DNA to *Cot1* DNA, 1:20), resuspended in 15 μl of hybridization solution (50% formamide, 2× SSC, 10% dextran sulphate, 0.15% SDS) and co-denatured on slides for 5 min at 76 °C. Slides were then incubated at 37 °C for 16 h, washed in 50% formamide/2× SSC for 5 min twice at 45 °C, then in 2× SSC for 5 min twice at 45 °C and incubated with avidin–Cy3 and antidigoxigenin–FITC (Roche, 1:250 dilution) in 2× SSC/0.05% Tween-20 for 1 h at room temperature (25 °C). After three washes in 2× SSC/0.05% Tween-20, slides were mounted in Vectashield with DAPI (Vector Laboratories). Metaphase images were captured using a Nikon Eclipse microscope equipped with a CCD camera (Applied Spectral Imaging) and a ×63 objective lens. Fifty metaphases were analysed for each sample.

Telomere-FISH. Metaphases on slides were prepared as described above. After pepsin digestion (1 mg ml⁻¹ for 10 min at 37 °C), slides were denatured at 80 °C for 3 min, hybridized with a Cy3-labelled PNA telomeric probe (Cy3-(TTAGGG)₃) in 70% formamide at room temperature for 2 h, washed, dehydrated and mounted in Vectashield with DAPI (Vector Laboratories). Metaphase images were captured using a Nikon Eclipse microscope equipped with a CCD camera (Applied Spectral Imaging) and a ×63 objective lens. For cytogenetic analysis, 30–40 metaphases per sample were scored for chromosomal aberrations.

32. Jones, K. R. *et al.* Radiosensitization of MDA-MB-231 breast tumor cells by adenovirus-mediated overexpression of a fragment of the XRCC4 protein. *Mol. Cancer Ther.* 4, 1541–1550 (2005).

Rag mutations reveal robust alternative end joining

Barbara Corneo^{1,2*}, Rebecca L. Wendland^{1*}, Ludovic Deriano^{1*}, Xiaoping Cui³, Isaac A. Klein¹, Serre-Yu Wong¹, Suzette Arnal¹, Abigail J. Holub¹, Geoffrey R. Weller¹, Bette A. Pancake¹, Sundeep Shah^{4†}, Vicky L. Brandt¹, Katheryn Meek³ & David B. Roth¹

Mammalian cells repair DNA double-strand breaks (DSBs) through either homologous recombination or non-homologous end joining (NHEJ). V(D)J recombination, a cut-and-paste mechanism for generating diversity in antigen receptors, relies on NHEJ for repairing DSBs introduced by the Rag1–Rag2 protein complex. Animals lacking any of the seven known NHEJ factors are therefore immunodeficient¹. Nevertheless, DSB repair is not eliminated entirely in these animals: evidence of a third mechanism, ‘alternative NHEJ’, appears in the form of extremely rare V(D)J junctions^{2–4} and a higher rate of chromosomal translocations^{5,6}. The paucity of these V(D)J events has suggested that alternative NHEJ contributes little to a cell’s overall repair capacity, being operative only (and inefficiently) when classical NHEJ fails. Here we find that removing certain portions of murine Rag proteins reveals robust alternative NHEJ activity in NHEJ-deficient cells and some alternative joining activity even in wild-type cells. We propose a two-tier model in which the Rag proteins collaborate with NHEJ factors to preserve genomic integrity during V(D)J recombination.

In previous work we discovered that the Rag proteins have an essential role in the joining phase of V(D)J recombination, but it has been unclear how the Rag post-cleavage complex transmitted the broken DNA ends specifically to the classical NHEJ pathway, rather than other repair machinery, for joining^{7–9}. Here we used a genetic approach, to be described elsewhere, to screen for Rag mutations that affect joining. We identified a mutant Rag2 protein with a frameshift at amino acid 361 (FS361), deleting the carboxy-terminal 22 amino acids of core Rag2 and adding 27 novel amino acids (Supplementary Fig. 2). Although core Rag2 has been considered the shortest active form of the protein^{10,11}, this FS361 mutant is fully active for DNA cleavage and recombination, as is an alternative version bearing a stop codon at the site of the frameshift (Stop361) (Supplementary Figs 2 and 3). The FS361 mutation did not affect steady-state levels of Rag2, as assessed by western blotting, but the Stop361 mutation yielded lower protein levels, indicating that it diminishes protein stability (Supplementary Fig. 4).

We sought to determine whether these Rag2 mutants allow V(D)J recombination intermediates (signal and coding ends) to access repair pathways other than classical NHEJ. We first used a substrate that allows Rag cleavage to generate a functional cyan fluorescent protein (CFP) gene through homologous recombination (Supplementary Fig. 5)⁹. FS361, Stop361 and core Rag2 all yielded statistically significant ($P < 0.003$) increases in CFP-positive cells compared with wild-type Rag2, which performs no better than the catalytically inactive ‘DDE’ mutant¹² (Fig. 1a). We next designed a recombination substrate that specifically measures alternative NHEJ,

making use of the knowledge that alternative NHEJ often produces large deletions and typically uses short sequence homologies^{1,5,13–15}. This substrate allows a functional green fluorescent protein (GFP)

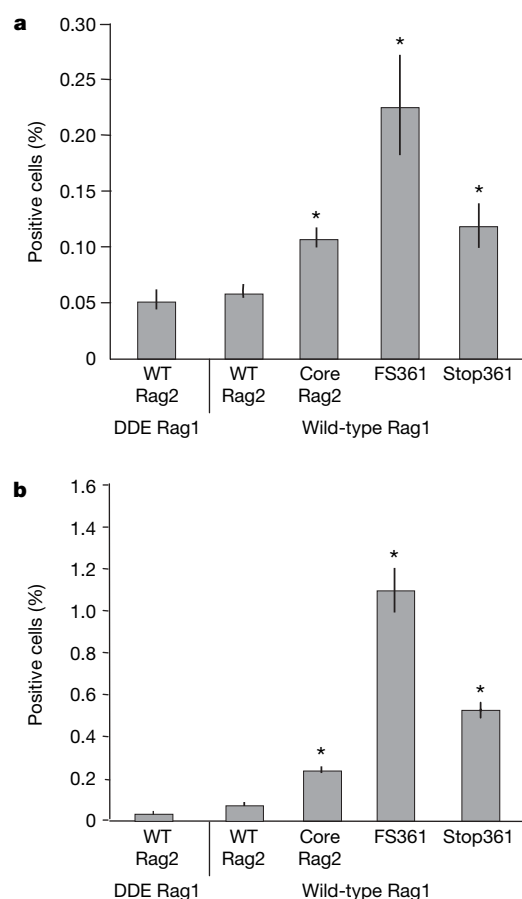


Figure 1 | FS361 and Stop361 Rag2 mutants allow use of homologous recombination and alternative NHEJ in wild-type cells. **a**, To measure homologous recombination in CHO-K1 cells, CFP⁺ fluorescent cells were measured by flow cytometry 60 h after transfection with the indicated Rag expression vectors and the substrate plasmid. **b**, To measure alternative NHEJ in CHO-K1 cells, GFP⁺ fluorescent cells were measured as in **a**. Values significantly different from the wild-type Rag1 and Rag2 complex are marked with an asterisk ($P < 0.05$ by Student’s *t*-test, unpaired, two tailed, equal variance). For **a** and **b**, means from three independent experiments are plotted; error bars reflect the standard error of the mean.

¹The Kimmel Center for Biology and Medicine of the Skirball Institute and Department of Pathology, New York University School of Medicine, New York, New York 10016, USA.

²Department of Gene and Cell Medicine, The Black Family Stem Cell Institute, Mount Sinai School of Medicine, New York, New York 10029, USA. ³College of Veterinary Medicine, Pathobiology & Diagnostic Investigation, Michigan State University, East Lansing, Michigan 48824, USA. ⁴Department of Immunology, Baylor College of Medicine, Houston, Texas 77030, USA. [†]Present address: Vertex Pharmaceuticals, 11010 Torreyana Road, San Diego, California 92121, USA.

*These authors contributed equally to this work.

gene to form only if a unique coding joint sequence is generated within a 9-nucleotide microhomology, necessitating deletion of 20 nucleotides from the coding ends (Supplementary Fig. 6). As expected, these highly restrictive conditions prevented the wild-type Rag complex from producing more than a few GFP-positive cells (Fig. 1b). FS361 and Stop361, however, generated much higher levels of GFP-positive cells (for FS361, a 12-fold increase over wild-type, $P < 0.0001$; for Stop361, a 6-fold increase, $P < 0.0001$) (Fig. 1b). Note that both of these increases are much larger than the slight increase in normal recombination (\leq twofold) observed with these two mutants (Supplementary Fig. 3). Core Rag2 modestly stimulated formation of alternative NHEJ products (threefold over wild-type; $P < 0.0001$) (Fig. 1b), as was the case with the homologous recombination substrate.

Given that removal of the Rag2 C terminus allows V(D)J recombination intermediates to be repaired by alternative NHEJ in wild-type cells, we wondered whether our mutants would also overcome the severe joining defects of NHEJ-deficient cells. We first tested FS361 using extrachromosomal substrates in V3 cells (a CHO cell line), which lack DNA-PKcs, the catalytic subunit of the NHEJ factor DNA-dependent protein kinase¹⁶. In the deficient cells, FS361 markedly and consistently ($P < 0.05$) rescued signal and coding joint formation (to 100% and 50%, respectively, of levels formed by wild-type Rag proteins in V3 cells complemented with DNA-PKcs) (Fig. 2a, b and Supplementary Table 1). To correct for the fact that coding and signal joint formation are both substantially stimulated by FS361 in the complemented cells, we expressed joint formation as a percentage of joints formed by the same Rag proteins in complemented cells. According to this calculation, FS361 formed 6% and 67% of wild-type levels of coding and signal joints, respectively (Fig. 2), a substantial rescue of joining.

We next examined coding and signal joint formation in a CHO cell line (XR-1) that lacks Xrcc4, a protein essential for classical NHEJ¹⁷. These are among the most severely V(D)J-recombination-deficient cells known. In mock-complemented XR-1 cells, rescue of signal

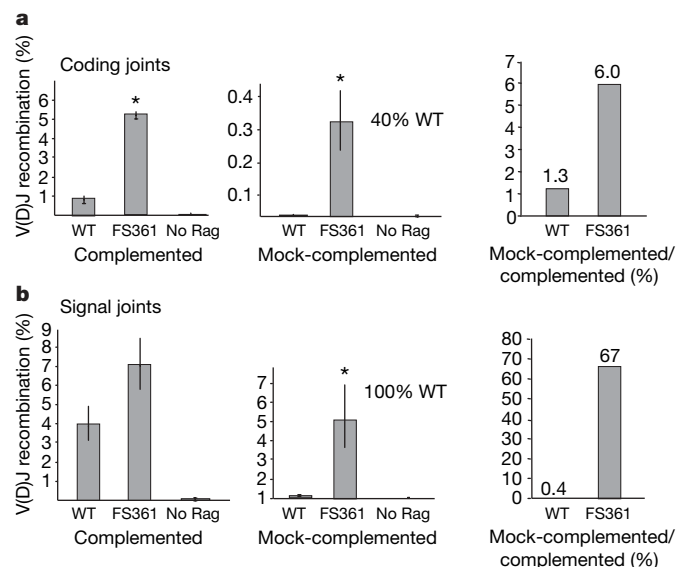


Figure 2 | Rag2 truncation mutants bypass genetic defects in classical NHEJ (DNA-PKcs). **a, b,** FS361 bypasses the NHEJ deficiency to form coding joints (at 40% the level of joint formation in wild-type cells) and rescues signal joint formation to the levels observed in wild-type cells. In the right panels data from the left and middle panels were used to calculate the ratio of recombination frequencies of mock-complemented to complemented cells, expressed as a percentage. This was done to correct for possible bias introduced by the varying abilities of the different Rag proteins to perform recombination. Data are from Supplementary Table 1. Asterisks indicate $P < 0.05$; error bars indicate standard error of the mean. (see Methods for details)

joint formation by FS361 was remarkably robust, averaging 40% of levels observed in complemented XR-1 cells transfected with wild-type Rag1 and Rag2 (Supplementary Fig. 7a and Supplementary Table 2a). Coding joint formation increased from undetectable levels to 10% of those observed with wild-type Rag proteins in complemented XR-1 cells (Supplementary Fig. 7a and Supplementary Table 2b).

To control for the higher levels of recombination observed with FS361 in complemented cells (Supplementary Fig. 7), we expressed the level of joining as a percentage of recombination observed in complemented XR-1 cells using the same Rag alleles (Fig. 3a). These calculations reveal that FS361 rescued signal joint formation in Xrcc4-deficient cells to 21% of levels in wild-type cells, strongly bypassing Xrcc4 deficiency. Stop361 and, to a lesser degree, core Rag2 also bypassed the signal joint defect in XR-1 cells (Fig. 3a). FS361 rescued coding joint formation to 3% of levels observed in the complemented control cells (Fig. 3a). This represents a 60-fold increase over the ratio of levels of coding joints produced by wild-type Rag2 in mock-complemented cells. Stop361 also rescued coding joint formation (sevenfold over the ratio of levels observed with wild-type Rag2). The weaker activity of Stop361 is consistent with the diminished stability of this protein, as noted above.

To investigate the effects of the dispensable portions of Rag1, we repeated the above experiments, replacing wild-type Rag1 with core Rag1 (refs 18, 19) (Supplementary Fig. 2). Core Rag1 amplified the ability of FS361 to bypass Xrcc4 deficiency, producing coding joints at 10% and signal joints at 38% of levels observed with the corresponding Rag proteins in wild-type cells (Fig. 3b). Notably, core Rag1 allowed some bypass even in the context of wild-type Rag2, forming signal joints at 6% of levels seen in wild-type cells (Fig. 3b).

Nucleotide sequences of coding joints obtained from Xrcc4-deficient cells transfected with FS361 resemble the rare junctions

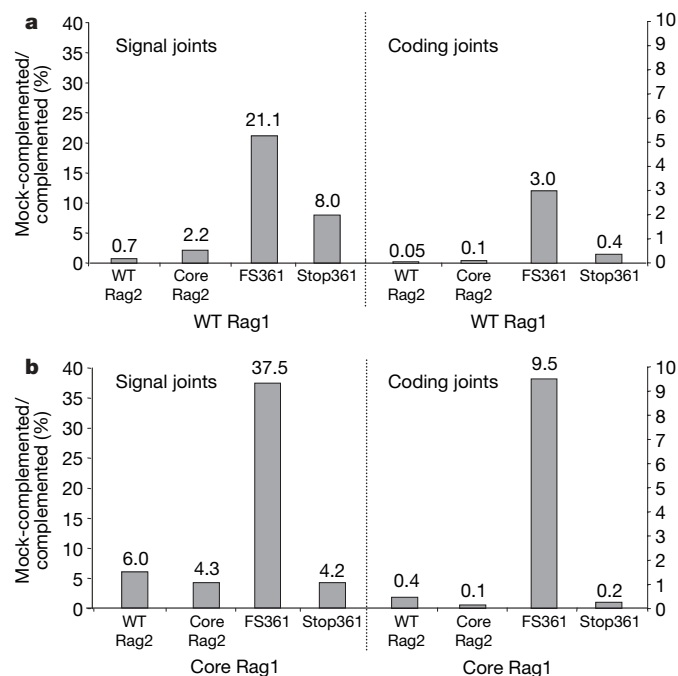


Figure 3 | Rag truncation mutants bypass genetic defects in classical NHEJ (Xrcc4). **a,** XR-1 cells complemented with an Xrcc4 expression vector were transfected with either a signal or coding joint substrate (pJH289 or pJH290, respectively) and the indicated Rag expression vectors. Plasmid DNA was harvested and used to transform bacteria; colonies were counted and used to calculate recombination frequencies (shown in Supplementary Table 2 and Supplementary Fig. 7). Results are expressed as a ratio of recombination frequencies of mock-complemented to complemented cells, as in Fig. 2, to correct for different recombination efficiencies of the Rag mutants. **b,** Experiments shown in **a** were repeated with core Rag1.

isolated from these cells using wild-type Rag proteins (Supplementary Table 3). Thus, this Rag2 truncation does not normalize NHEJ but rather allows the ends to access alternative NHEJ pathways. Notably, most coding joints (82%) exhibited microhomologies characteristic of alternative NHEJ (Supplementary Table 3). Indeed, 50% used a seven-nucleotide microhomology with a 35-base-pair deletion never detected in wild-type cells. Signal joints, which are normally precise, were strikingly abnormal: all except one (58 out of 59) from NHEJ-deficient cells bore deletions and 83% showed microhomologies of at least 3 nucleotides (Supplementary Table 3).

To test the ability of Rag mutants to bypass classical NHEJ defects under more physiological conditions, we repeated the aforementioned experiments using *Xrcc4*-deficient cells that contained chromosomally integrated coding joint and signal joint substrates. Polymerase chain reaction (PCR) analysis of genomic DNA revealed that Stop361 and FS361 were active for signal and coding joint formation in wild-type cells bearing integrated substrates (Supplementary Fig. 8) and rescued coding joint formation in *Xrcc4*-deficient cell lines (Fig. 4). In agreement with results using extrachromosomal substrates, the most robust rescue of coding joint formation occurred

with FS361 in combination with core Rag1, and both FS361 and Stop361 rescued coding joint formation in conjunction with full-length Rag1 (Fig. 4a). FS361 also strongly rescued signal joint formation, whether in combination with core or wild-type Rag1 (Fig. 4b). Core Rag1 in combination with wild-type Rag2 allowed a mild rescue of both coding and signal joint formation, consistent with the results obtained using extrachromosomal substrates.

The bypass observed in both *Xrcc4*- and DNA-PKcs-deficient cell lines indicates a much more robust level of alternative joining activity than previously appreciated. Our data suggest that classical NHEJ factors and the Rag post-cleavage complex collaborate to provide two tiers of protection against aberrant rearrangements (Supplementary Fig. 1). Recent data showing that immunoglobulin class-switch recombination is quite efficient in *Xrcc4*- and DNA-ligase-IV-deficient contexts²⁰ corroborate our conclusion that alternative NHEJ is quite active in NHEJ-deficient cells. Class switching requires proteins not necessary for V(D)J recombination²¹; perhaps these factors afford a second tier of protection against aberrant rearrangements in the absence of a class-switch recombination complex analogous to the Rag post-cleavage complex.

How active is alternative NHEJ in wild-type cells? Previous analysis of joints formed by re-circularization of transfected linear DNA molecules revealed a substantial fraction of junctions with deletions and short sequence homologies¹³. Using the substrate designed to measure alternative end joining, we found a substantial increase in the formation of a particular highly deleted junction bearing a long microhomology in the presence of FS361 and, to a lesser extent, Stop361 and core Rag2. These data suggest that deletions of the C terminus of Rag2 allow alternative NHEJ activity in wild-type cells. Sequences of coding and signal joints suggested a slight trend towards larger deletions but failed to show consistent, gross abnormalities typical of joints formed in NHEJ-deficient cells (data not shown). It is conceivable, however, that the presence of a functional NHEJ system might alter the features of junctions formed by the alternative pathway, perhaps by protecting the ends from excessive degradation. Indeed, even in *Xrcc4*-deficient cells a significant fraction of junctions appears normal, indicating that deletions and microhomologies are not obligatory features of alternative NHEJ.

The previously puzzling observation that core Rag1 and core Rag2 knock-in mice form coding and signal joints bearing excessive deletions and short sequence homologies²² can now be explained by our data showing that both core Rag1 and core Rag2 allow a small degree of alternative NHEJ activity. The presence of these aberrant junctions in core Rag1 and Rag2 knock-in mice also provides strong support for our conclusion that wild-type cells express alternative NHEJ activity. We speculate that preventing V(D)J recombination intermediates from being joined by alternative NHEJ benefits the organism by ensuring a diverse immune receptor repertoire (which would be severely constrained by frequent use of microhomologies and the excessive deletions characteristic of alternative NHEJ) and by minimizing the opportunities for chromosome translocations.

A shift in the balance between classical and alternative NHEJ could represent an important step in carcinogenesis^{23,24}. Biochemical studies have shown that human bladder tumours primarily express alternative NHEJ activity, as opposed to normal human cells that perform classical NHEJ²³. Given the unexpected robustness of the alternative NHEJ pathway, as shown by our studies and those of ref. 20, it is clearly important to determine the contributions of alternative NHEJ mechanisms to normal cell physiology as well as to the development of neoplasia.

METHODS SUMMARY

The FS361 mutant was generated by error-prone PCR-based mutagenesis of the *Rag2* open reading frame. The Stop361 mutant was created by site-directed mutagenesis. Transient transfections with fluorescent recombination reporter substrates were performed using FuGENE 6 (Roche) according to the manufacturer's directions. Extrachromosomal recombination assays using the bacterial

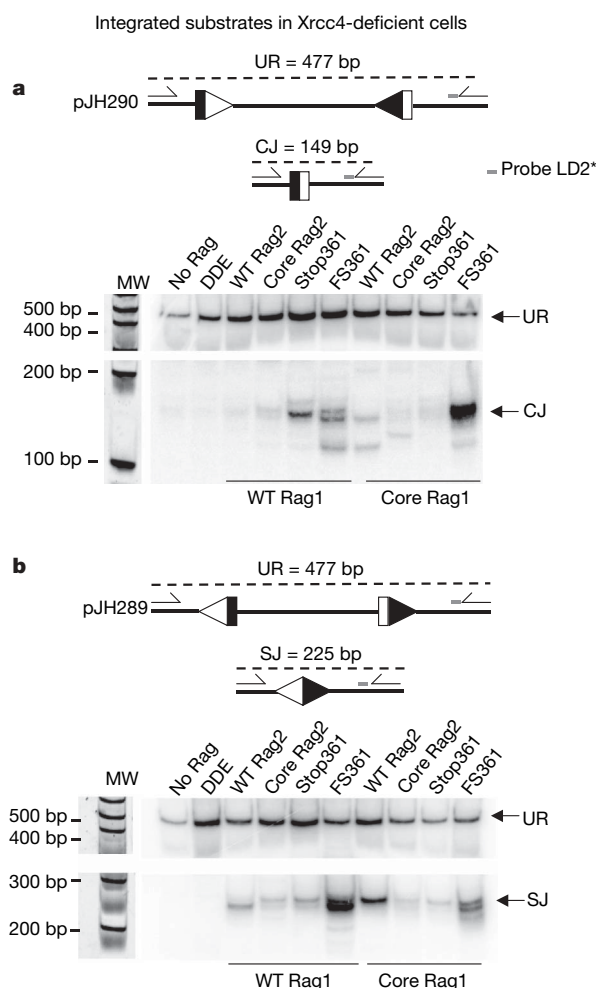


Figure 4 | FS361 and Stop361 bypass *Xrcc4* deficiency using integrated substrates. PCR analysis of genomic DNA harvested from integrated recombination substrates in *Xrcc4*-deficient cells. PCR products were run on a gel, blotted and probed with the indicated oligonucleotide probes. **a**, The most robust rescue of coding joint formation occurred with FS361 in combination with core Rag1. **b**, FS361 also strongly rescued signal joint formation, whether in combination with core or wild-type Rag1. Core Rag1 in combination with wild-type Rag2 allowed a mild rescue of both coding and signal joint formation, consistent with the results obtained with extrachromosomal substrates. CJ, coding joint; SJ, signal joint; UR, unrearranged substrate.

colony counting method were performed as described previously¹⁶. Alternative NHEJ was measured using an extrachromosomal substrate that gives GFP expression only if 10 nucleotides of extraneous DNA are removed from each end and joining occurs in a 9-nucleotide microhomology (see Methods). Homologous recombination was measured using an extrachromosomal reporter substrate as described previously⁹. Bulk populations of cells bearing stably integrated recombination substrates were maintained under selection for the appropriate resistance marker, except during recombination assays, and were tested for recombination by transiently transfecting Rag expression vectors. *P* values were calculated according to the Student's *t*-test (unpaired, two tailed, equal variance).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 June; accepted 17 August 2007.

- Roth, D. B. Restraining the V(D)J recombinase. *Nature Rev. Immunol.* **3**, 656–666 (2003).
- Schuler, W. *et al.* Rearrangement of antigen receptor genes is defective in mice with severe combined immune deficiency. *Cell* **46**, 963–972 (1986).
- Bogue, M. A., Wang, C., Zhu, C. & Roth, D. B. V(D)J recombination in Ku86-deficient mice: Distinct effects on coding, signal, and hybrid joint formation. *Immunity* **7**, 37–47 (1997).
- Gao, Y. *et al.* A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* **95**, 891–902 (1998).
- Zhu, C. *et al.* Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell* **109**, 811–821 (2002).
- Ferguson, D. O. & Alt, F. W. DNA double strand break repair and chromosomal translocation: lessons from animal models. *Oncogene* **20**, 5572–5579 (2001).
- Qiu, J. X., Kale, S. B., Yarnall Schultz, H. & Roth, D. B. Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. *Mol. Cell* **7**, 77–87 (2001).
- Yarnall Schultz, H., Landree, M. A., Qiu, J. X., Kale, S. B. & Roth, D. B. Joining-deficient RAG1 mutants block V(D)J recombination *in vivo* and hairpin opening *in vitro*. *Mol. Cell* **7**, 65–75 (2001).
- Lee, G. S., Neiditch, M. B., Salus, S. S. & Roth, D. B. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* **117**, 171–184 (2004).
- Cuomo, C. A. & Oettinger, M. A. Analysis of regions of RAG-2 important for V(D)J recombination. *Nucleic Acids Res.* **22**, 1810–1814 (1994).
- Sadofsky, M. J., Hesse, J. E. & Gellert, M. Definition of a core region of RAG-2 that is functional in V(D)J recombination. *Nucleic Acids Res.* **22**, 1805–1809 (1994).
- Landree, M. A., Wibbenmeyer, J. A. & Roth, D. B. Mutational analysis of RAG-1 and RAG-2 identifies three active site amino acids in RAG-1 critical for both cleavage steps of V(D)J recombination. *Genes Dev.* **13**, 3059–3069 (1999).
- Roth, D. B. & Wilson, J. H. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Mol. Cell. Biol.* **6**, 4295–4304 (1986).
- Kabotyanski, E. B., Gomelsky, L., Han, J.-O., Stamato, T. D. & Roth, D. B. Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res.* **26**, 5333–5342 (1998).
- Verkaik, N. S. *et al.* Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. *Eur. J. Immunol.* **32**, 701–709 (2002).
- Ding, Q. *et al.* Autophosphorylation of the catalytic subunit of the DNA-dependent protein kinase is required for efficient end processing during DNA double-strand break repair. *Mol. Cell. Biol.* **23**, 5836–5848 (2003).
- Li, Z. *et al.* The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell* **83**, 1079–1089 (1995).
- Sadofsky, M. J., Hesse, J. E., McBlane, J. F. & Gellert, M. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Res.* **21**, 5644–5650 (1993).
- Kirch, S. A., Sudarsanam, P. & Oettinger, M. A. Regions of RAG1 protein critical for V(D)J recombination. *Eur. J. Immunol.* **26**, 886–891 (1996).
- Yan, C. T. *et al.* IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* doi:10.1038/nature06020 (this issue); published online 22 August 2007.
- Posey, J. E., Brandt, V. L. & Roth, D. B. Paradigm switching in the germinal center. *Nature Immunol.* **5**, 476–477 (2004).
- Talukder, S. R., Dudley, D. D., Alt, F. W., Takahama, Y. & Akamatsu, Y. Increased frequency of aberrant V(D)J recombination products in core RAG-expressing mice. *Nucleic Acids Res.* **32**, 4539–4549 (2004).
- Bentley, J., Diggle, C. P., Harnden, P., Knowles, M. A. & Kiltie, A. E. DNA double strand break repair in human bladder cancer is error prone and involves microhomology-associated end-joining. *Nucleic Acids Res.* **32**, 5249–5259 (2004).
- Deriano, L., Merle-Beral, H., Guipaud, O., Sabatier, L. & Delic, J. Mutagenicity of non-homologous end joining DNA repair in a resistant subset of human chronic lymphocytic leukaemia B cells. *Br. J. Haematol.* **133**, 520–525 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by NIH grants (to K.M. and D.B.R.) and the Irene Diamond Fund (D.B.R.). L.D. is a Fellow of The Leukemia and Lymphoma Society and G.R.W. is a Cancer Research Institute Postdoctoral Fellow. We thank J. Qiu for help with the mutant library and S. Salus for help during the early phases of this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.B.R. (roth@saturn.med.nyu.edu).

METHODS

Cell culture. Chinese hamster ovary (CHO-K1 and RMP41) fibroblasts and the CHO mutant cell lines XR-1 (deficient in *Xrcc4*) and V3 (deficient in DNA-PKcs) were grown in DMEM (Invitrogen) supplemented with fetal bovine serum (10%), non-essential amino acids and penicillin-streptomycin. Cells were grown at 37 °C in the presence of 5% CO₂. Cells containing integrated substrates were selected and grown in the presence of 250 µg ml⁻¹ geneticin (Gibco).

Measuring V(D)J recombination by bacterial colony analysis. To assess V(D)J recombination in XR-1 cells, 1.8 µg of murine Rag1 and 2.1 µg of murine Rag2 expression constructs, as well as 5 µg of recombination substrate pJH289 (ref. 23; signal joints) or pJH290 (ref. 23; coding joints), were transfected into cells stably expressing murine *Xrcc4* or mock vector using a FuGENE 6:DNA ratio of 3:1. Forty-eight hours after transfection, plasmid substrates were rescued from the cells and bacteria transformed with these DNAs were spread onto LB plates containing 100 µg ml⁻¹ ampicillin and LB plates containing 100 µg ml⁻¹ ampicillin and 11 µg ml⁻¹ chloramphenicol. A parallel approach was used for the studies with V3 (DNA-PKcs-deficient) cells.

Fluorescent reporter substrates to detect V(D)J coding and signal joint formation. We inserted the *Bgl*II to *Eco*RI recombination signal sequence (RSS)-bound segment from pJH290 into pGFP-N2 (Clontech) to generate plasmid pSS1. Subsequently we inserted the *Sal*I to *Bam*HI segment containing a poly(A) sequence from plasmid dsRed (Clontech) between the 12 and 23 RSS to generate pCJGFP to detect coding joint formation. Before Rag cleavage, GFP expression is blocked by the poly(A) sequence. The signal-joint-specific reporter pSJGFP was constructed similarly, except the 12/23 RSS are inverted relative to the RSS sequences in pJH290.

To measure V(D)J recombination, cells were grown to 50% confluence in 24-well dishes. Each transfection contained 400–600 ng of each plasmid, substrate as indicated, and Rag expression vectors as indicated, and a FuGENE 6 (Roche):DNA ratio of 3:1 and 3:2 for CHO-K1 and RMP41, respectively.

Flow cytometry. FACS analysis was carried out using a BD LSRII flow cytometer (BD Biosciences) equipped with FACS Diva software after cells were trypsinized, pelleted and re-suspended in DPBS (Invitrogen) containing 0.5% BSA (Sigma Aldrich) and 5 mM EDTA. FlowJo software was used for data analysis.

Substrate-specific integrated cell lines. 3 × 10⁵ RMP41 cells were plated and transfected 48 h later with 800–1,000 ng of either the pCJGFP (coding joint) or pSJGFP (signal joint) substrate using the FuGENE 6 transfection reagent (Roche) at a 3:2 FuGENE:DNA (µl:µg) ratio. Integrated cells were selected with 250 µg ml⁻¹ of geneticin (Gibco/Invitrogen) for 14 days. Bulk, stably integrated cell lines were maintained under antibiotic selection except during recombination assays.

1 × 10⁶ CHO-K1 and XR-1 cells were plated and transfected 24 h later with 6 µg of either the pJH290 (coding joint) or pJH289 (signal joint) substrate²⁵ and

with 2 µg of pBABEpuro plasmid (ADDGENE) using the FuGENE 6 transfection reagent (Roche) at a 3:1 FuGENE:DNA (µl:µg) ratio. Integrated cells were selected with 3 µg ml⁻¹ of puromycin for 12 days. Bulk, stably integrated cell lines were maintained under antibiotic selection except during recombination assays.

Recombination assays in pJH290 and pJH289 integrated cell lines. To assess V(D)J recombination in substrate-specific integrated XR-1 and CHO-K1 cell lines, 2 µg of Rag1 and 2 µg of Rag2 expression constructs were transfected into cells using a FuGENE 6:DNA ratio of 3:1. Forty-eight hours after transfection, genomic DNA was extracted and substrate-specific V(D)J recombination events were analysed by PCR using the substrate-specific primers DR99 and DR100. PCR products were analysed by polyacrylamide gel electrophoresis and visualized using a PhosphorImager after Vistra Green staining (GE Healthcare Biosciences) or Southern blotting using a ³²P-end-labelled oligo probe, LD2.

Primer and probe sequences. DR99, 5'-TCACACAGGAAACAGCTATGAC-CATG-3'; DR100, 5'-GGGATATATCAACGGTGGTATATCCAGTG-3'; LD2, 5'-GAGATTTTCAGGAGCTAAGGAAGCTAAAAATGG-3'.

Alternative NHEJ assays. Cells were grown to 50% confluence in 24-well dishes and transfected with 400 ng each of Rag1, Rag2 and the GFP alternative NHEJ substrate, pGFP-Alt, using a FuGENE 6:DNA ratio of 3:1. Cells were harvested at 48 h and analysed by flow cytometry. The alternative NHEJ substrate was created by adding 10 nucleotides of extraneous DNA (with stop codons in all three reading frames) and 9 nucleotides of DNA homologous to the interior of the coding region for pEGFP-N1 (Clontech) to either side of the heptamer-bound segment from pEGFP-290-NtCFP⁹ by PCR. This segment was then inserted to interrupt the EGFP coding sequence of pEGFP-N1, yielding pGFP-Alt. Formation of a functional GFP gene requires cleavage, deletion of 10 nucleotides from each end and utilization of the 9-nucleotide microhomology.

Homologous recombination. Assays were performed using an extrachromosomal reporter substrate that expresses CFP on homologous recombination. Cells were harvested at ~60 h after transfection and analysed by flow cytometry to detect CFP expression. We considered any cells brighter than substrate alone (without Rag vectors) to be positive.

Western blot analysis. After transient transfections of RMP41 cells, proteins from cell lysates were separated on 8% SDS-PAGE gels and transferred to PVDF membranes (Bio-Rad). Rag protein expression was detected by chemiluminescence using rabbit anti-GST (Cell Signaling Technology) and donkey anti-rabbit IgG conjugated with HRP (GE Healthcare Biosciences) and visualized using the ECL-PLUS kit (GE Healthcare Biosciences).

25. Lieber, M. R. *et al.* The defect in murine severe combined immune deficiency: joining of signal sequences but not coding segments in V(D)J recombination. *Cell* 55, 7–16 (1988).

Structural basis of Dscam isoform specificity

Rob Meijers^{1,3,*†}, Roland Puettmann-Holgado^{2,4,8*}, Georgios Skiniotis⁵, Jin-huan Liu^{1,3}, Thomas Walz⁵, Jia-huai Wang^{1,6,7} & Dietmar Schmucker^{2,4}

The *Dscam* gene gives rise to thousands of diverse cell surface receptors¹ thought to provide homophilic and heterophilic recognition specificity for neuronal wiring^{2–4} and immune responses⁵. Mutually exclusive splicing allows for the generation of sequence variability in three immunoglobulin ecto-domains, D2, D3 and D7. We report X-ray structures of the amino-terminal four immunoglobulin domains (D1–D4) of two distinct *Dscam* isoforms. The structures reveal a horseshoe configuration, with variable residues of D2 and D3 constituting two independent surface epitopes on either side of the receptor. Both isoforms engage in homo-dimerization coupling variable domain D2 with D2, and D3 with D3. These interactions involve symmetric, antiparallel pairing of identical peptide segments from epitope I that are unique to each isoform. Structure-guided mutagenesis and swapping of peptide segments confirm that epitope I, but not epitope II, confers homophilic binding specificity of full-length *Dscam* receptors. Phylogenetic analysis shows strong selection of matching peptide sequences only for epitope I. We propose that peptide complementarity of variable residues in epitope I of *Dscam* is essential for homophilic binding specificity.

The *Drosophila melanogaster* *Dscam* protein is an immunoglobulin superfamily (IgSF) member and consists of 10 immunoglobulin-like domains, 6 type III fibronectin domains, a transmembrane segment and a 374-residue cytoplasmic domain (Supplementary Fig. 1). Mutually exclusive alternative splicing occurs for exons 4, 6, 9 and 17, which encode the N-terminal half of Ig2 (D2), the N-terminal half of D3, the entire D7, and the transmembrane segment, respectively. In *Drosophila*, the combinatorial use of alternative exons potentially gives rise to 19,008 distinct extracellular receptor parts of *Dscam*¹. *Dscam* was first identified as a highly diverse surface receptor required for neuronal wiring^{1,6–13}. It was proposed that differences in isoforms expressed on the surface of neighbouring axons and dendrites determine their interactions^{2,4,8}. Recent studies have demonstrated that *Dscam* is also required in the innate immune system^{5,14}. Moreover, *in vitro* binding studies have shown that *Dscam* isoforms can interact in a highly selective homophilic manner, and that even closely related isoforms exhibit almost exclusive isoform-specific binding requiring the first eight immunoglobulin domains². It was also proposed that modular interactions of the variable domains as D2^A–D2^B, D3^A–D3^B and D7^A–D7^B together are required to stabilize otherwise weak individual interactions².

The topology of the N-terminal eight immunoglobulin domains of *Dscam* was examined by negative-stain electron microscopy¹⁵. Averaged images of several isoforms reveal multiple distinct configurations indicative of remarkable flexibility (Fig. 1a, and Supplementary Fig. 2). In contrast, class averages of proteins of the N-terminal four immunoglobulin domains (D1–D4) revealed a horseshoe shape

for more than 90% of the molecules observed (Fig. 1b, and Supplementary Fig. 3).

D1–D4 proteins of two distinct *Dscam* isoforms were expressed using a baculovirus system and taken for crystallographic analysis (see Methods). The first isoform contains splicing variant 1 of exon 4 and splicing variant 34 of exon 6, designated as D1–D4_{1,34}. It crystallized in three forms, the best of which, *P*₄22, diffracted to 2.0 Å resolution. The 388-residue structure was determined using multiple anomalous dispersion and crystal averaging (Supplementary Table 1). It reveals a horseshoe-shaped configuration with D2 contacting D3 on the top and D1 contacting D4 at the bottom (Fig. 1c). The second isoform D1–D4_{9,9} contains splicing variant 9 of exon 4 and variant 9 of exon 6. Its 3 Å resolution 391-residue structure in C2 form was obtained by molecular replacement. In spite of marked sequence differences in D2 and D3 (Supplementary Fig. 4), D1–D4_{9,9} has a similar horseshoe configuration. Unless otherwise stated, D1–D4_{1,34} is used for a description of structural features.

The horseshoe configuration is facilitated by a flexible 5-residue linker between D2 and D3 (Fig. 1c) and buries 2,550 Å² of surface area in the *Dscam* D1–D4_{1,34} structure, a size that probably renders the configuration stable¹⁶. The interactions that stabilize the D2–D3 interface are mainly from conserved residues (Supplementary Fig. 5a). Because D1 and D4 are constant domains, the D1–D4 interface is also probably conserved among all isoforms. Therefore the horseshoe configuration should be a general structural feature of all *Dscam* isoforms. This overall horseshoe topology is similar to the N-terminal four-domain structures of the insect protein hemolin¹⁷ and the chicken neuronal receptor axonin¹⁸ (Supplementary Figs 5b and 6). The prediction is that this configuration might represent a common feature for other neural cell adhesion molecules of the L1 family¹⁹.

The four *Dscam* N-terminal domains most closely resemble the I-set of the IgSF fold²⁰. The alternatively spliced segment of D2 (Supplementary Fig. 4) exhibits two features that are unique to *Dscam*. First, both isoforms contain an unusually long A' strand (Fig. 1c), and, second, an extra β-strand (termed D⁰) that emerges after the C' strand, folding alongside the D strand (Fig. 1c, and Supplementary Fig. 7a, b). This C'–D⁰ unit encompasses the most variable region of exon 4 (Supplementary Fig. 4). The alternatively spliced segment in D3 contains an unusually elongated transition of the A–A'. In D1–D4_{1,34}, this prominent 8-residue A–A' bulges out (Fig. 1c, and Supplementary 7c), whereas in the D1–D4_{9,9}, this 7-residue A–A' curves into a single 3₁₀ helical turn (Supplementary Fig. 7d). The long A' strand of D2 and the distinct A–A' protrusion of D3 are exposed on the same front face and, as a consequence of the horseshoe configuration, assemble into a composite surface epitope. It displays variable residues that we define and refer to as epitope I

¹Department of Medical Oncology and ²Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ³Department of Medicine, ⁴Department of Neurobiology, ⁵Department of Cell Biology, ⁶Department of Pediatrics and ⁷Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁸Institute for Biology III, University of Freiburg, D-79104 Freiburg im Breisgau, Germany. †Present address: Synchrotron Soleil, L'Orme des Merisiers, 91192 Saint Aubin, France.

*These authors contributed equally to this work.

(Fig. 1d, e). The C'-D⁰ unit of D2 and the β -strand B of D3 reside on the opposite face, constituting an entirely separate epitope, which we define and refer to as epitope II (Fig. 1d, e). Both exon 4 and exon 6 contribute to each epitope in such a manner that their N-terminal segments encode epitope I and the C-terminal segments encode epitope II sequences.

Molecular packing analyses of all crystal forms of both isoforms show that epitope I is the key interaction surface engaged in

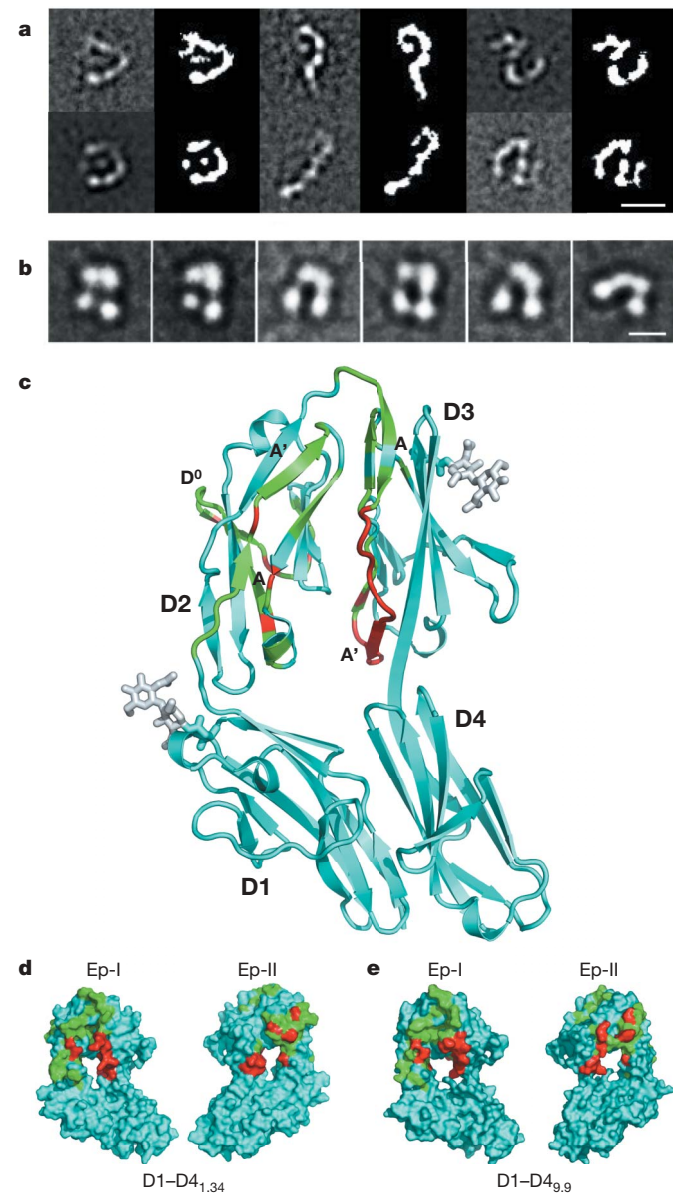


Figure 1 | Structure of the N-terminal four-domain fragment of Dscam.

a, Representative class averages from negatively stained Dscam D1-D8_{1,34,30} show that the molecule can adopt different conformations but retains the horseshoe configuration of the N-terminal D1-D4 domains. Scale bar, 10 nm. **b**, Representative class averages from negatively stained Dscam D1-D4_{1,34} show that the four domains of Dscam D1-D4_{1,34} are arranged in a horseshoe configuration. Scale bar, 5 nm. **c**, Ribbon diagram of Dscam D1-D4_{1,34} coloured according to sequence variability; conserved residues are coloured cyan, variable residues are green and hypervariable residues are red. The variability was calculated using Shannon's uncertainty²², and residues were classified as hypervariable if the uncertainty value exceeded two-thirds of the highest value observed for all residues from exons 4 and 6. **d**, **e**, Surface representation of epitope I (left) and II (right) on either side of the horseshoe for Dscam D1-D4_{1,34} (**d**) and Dscam D1-D4_{9,9} (**e**). Colour codes are as in **c**. The figure was prepared using PyMOL (<http://www.pymol.org>).

homophilic dimerization (Fig. 2). The central structural element in the D2^A-D2^B interactions between the monomers A and B is the long and highly variable A' strand. In D1-D4_{1,34}, this segment is EADVNKEH, starting at residue 107. A hydrogen-bonding network is formed between peptide segments Glu 107 to His 114 of interacting D2s (Fig. 2a, c). At the centre of the dimer are two side-chain hydrogen bonds formed by a pair of symmetry-related residues (Asn 111). There are two salt bridges (Glu 107^A-Lys 112^B and Asp 109^A-Lys 112^B) as well as a hydrogen bond (Asp 109^A-His 114^B) from molecules A to B, and vice versa. In D1-D4_{9,9}, the equivalent segment, starting at residue 109, is ESEADNEY. The phenol rings of dyad-related Tyr 114 pack against each other (Fig. 2b, d) at the centre of the D2^A-D2^B interface. The Tyr 114 side chain of one molecule is anchored by a hydrogen bond with the side chain of the conserved residue Thr 206 of a dyad-related molecule.

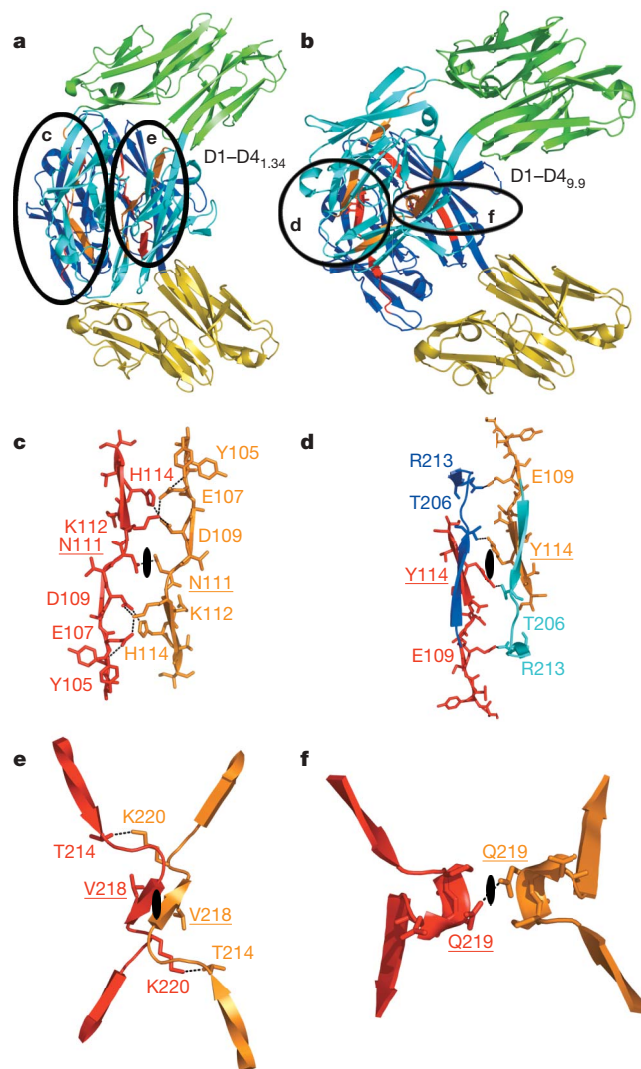


Figure 2 | Homophilic dimers observed in the crystal lattice. **a**, **b**, Ribbon diagram of the dimer in Dscam D1-D4_{1,34} (**a**) and Dscam D1-D4_{9,9} (**b**). D1 and D4, green; D2 and D3, blue for monomer A; D1 and D4, yellow; D2 and D3, cyan for monomer B. Residues at symmetry centre are underlined. The isoform-specific interaction elements are shown as red and orange in molecules A and B, respectively, and are displayed in more detail along their respective twofold axes: **c**, the D2^A-D2^B interface of Dscam D1-D4_{1,34}; **d**, the D2^A-D2^B interface of Dscam D1-D4_{9,9} (blue and cyan residues are constant); **e**, the D3^A-D3^B interface of Dscam D1-D4_{1,34}; **f**, the D3^A-D3^B interface of Dscam D1-D4_{9,9}. Residues involved in dimer-sustaining hydrogen bonds are labelled and the dyad axes are displayed as black ellipsoids. The figure was prepared using PyMOL (<http://www.pymol.org>).

The central structural element in the D3^A–D3^B interface is the highly variable A–A' protrusion. In D1–D4_{1,34}, the isoform-specific sequence for this segment, starting at residue 214, is TPALVQKP. From Ala 216 to Lys 220, the main chains of the two protomers pair into an anti-parallel mini β -sheet. The interaction is further stabilized by a symmetry-pair of specific hydrogen bonds (Lys 220^A–Thr 214^B). A pseudo-dyad runs through the middle of this β -sheet, where two symmetry-related Val 218 residues touch (Fig. 2e). In the D1–D4_{9,9} dimer, there are, again, interactions between these opposing A–A' protrusions, with a sequence of VNPQDKH starting at residue 216, but of a different nature. The most exposed residues Pro 218 and Gln 219 at the 3₁₀ helical turn's acme contribute prominent contacts to the dimer interface (Fig. 2f). One hydrogen bond is formed between side chains of a symmetry-related pair of Gln 219. In both isoforms, we noticed only minor cross interactions between D2 and D3, namely a salt bridge between the variable Arg 213 of D3 and the conserved Asp 133 of D2 in D1–D4_{1,34}, and a similar pair between Arg 213 of D3 and Glu 109 of D2 (both variable residues) in D1–D4_{9,9}.

Three key features characterize the homophilic dimerization of both isoforms: (1) the most variable residues of epitope I but not epitope II constitute the core of the dimer interface; (2) in a similar manner to other immunoglobulin-domain-containing adhesion molecules²¹, the interface is dominated by hydrophilic residues, and the binding strength is expected to be low; and (3) within the dimer there are separable D2^A–D2^B and D3^A–D3^B interfaces between molecules A and B. These shared features are predicted to be applicable to all 576 (12 D2 \times 48 D3) potential homo-dimer interactions. Although the central contact area of dimerization is clearly defined by the variable residues of epitope I in the two isoform structures, their orientation and the significantly different extent of buried surface area (3,711 Å² for the D1–D4_{1,34} dimer and 1,220 Å² for the D1–D4_{9,9} dimer) indicate a considerable diversity in the orientation and binding strength of homophilic isoform pairs.

Biochemical experiments indicate that Dscam homophilic binding is remarkably robust. However, homophilic interactions were only observed if the isoforms are nearly identical and contain the first eight immunoglobulin domains². Therefore, the question arises as to how much the local intermolecular contacts of D2 and D3 in D1–D4 homodimers (Fig. 2) contribute to the specificity of full-length Dscam homodimers. To address this, we introduced point mutations in residues critical to D2^A–D2^B and D3^A–D3^B interactions. Homophilic binding was tested using a previously described bead-to-cell adhesion assay². Beads that contain the entire extracellular part of the Dscam_{1,34,30} isoform bind efficiently to COS cells expressing the identical Dscam_{1,34,30} receptor but not at all to cells expressing Dscam_{1,30,30} (Fig. 3a, and Supplementary Figs 8 and 9).

In D1–D4_{1,34}, the D2^A–D2^B interface consists of the A' strand from residues Glu 107 to His 114. Therefore, K112E and H114D mutations will probably disrupt the hydrogen bonding in the interface. Indeed, the bead-to-cell binding assay shows that this double mutation reduces homophilic binding to only 4% compared to wild type (Fig. 3). The D3^A–D3^B interface is marked by the formation of a mini anti-parallel β -sheet, involving four main chain hydrogen bonds centred at Val 218. To disrupt this mini β -sheet, we mutated Val 218 to a proline, which reduced the homophilic binding to only 14% (Fig. 3a, and Supplementary Fig. 8).

Cell adhesion often relies on highly specific multivalent interactions, composed of relatively weak single interaction pairs²¹, which also seems to be the case for homophilic binding of Dscam. We mutated two residues within epitope I in tandem to alanine to reduce specific inter-domain hydrogen bonds. These mutations had a moderate effect on the cell adhesion properties of Dscam, reducing binding to 68% for N111A/K112A of D2 and 74% for L217A/Q219A of D3 (Fig. 3a, and Supplementary Fig. 8). When we combine tandem mutants for both domains, the reduction in adhesion is much more pronounced, with 31% for N111A/K112A/K220A/P221A and 10%

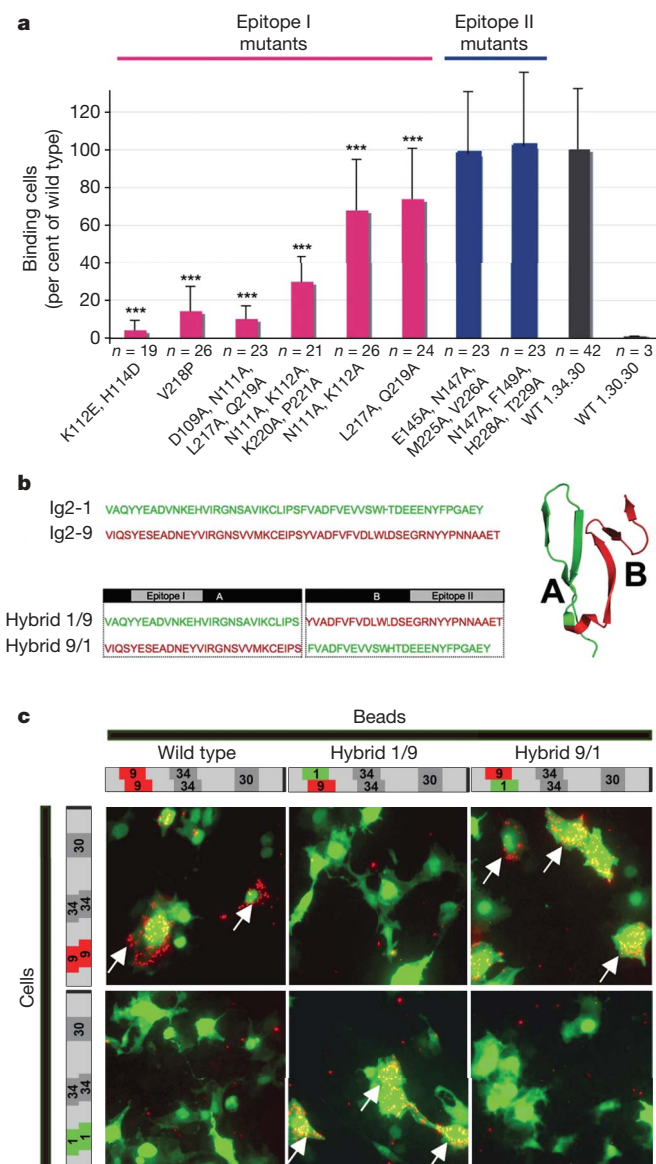


Figure 3 | Epitope I confers homophilic binding specificity.

a, Quantification of binding of wild-type Dscam receptor and eight different isoform variants bearing mutations in either epitope I or epitope II to Cos cells expressing wild-type Dscam_{1,34,30}. Representative micrographs depicting bead-to-cell binding are given in Supplementary Figs 9 and 10. The total numbers of counted fields, each containing 50–200 cells, are given below each bar. Error bars indicate standard deviation (as variation between entire optical fields (10 \times 10 microscope magnification)). *** P > 0.005, t -test. **b**, Peptide sequences for the two different Ig2 domains containing exon 4.1 or 4.9 (Ig2-1 or Ig2-9, respectively). Hybrid isoforms 1/9 and 9/1 were generated by overlap-PCR from exon 4 templates. Hybrid 1/9 incorporates the N-terminal half of exon 4 (segment A), including epitope I sequences from isoform 4.1, and the C-terminal half of exon 4 (segment B), including epitope II sequences from isoform 4.9. Hybrid 9/1 incorporates the equivalent 4.9 and 4.1 sequences. The structural model shows the distinct spatial separation of peptide segments. **c**, Bead-to-cell binding assays were performed (see Supplementary Figs 9 and 10) using Dscam protein with either wild-type Ig2-9 and Ig2-1 sequences (left panels) or hybrid sequences (middle and right panels for hybrid design see **b**). The schematics next to the micrographs indicate exon combinations (exon 4 in red or green boxes) of the proteins tested. Both hybrid proteins fully retained their homophilic binding capabilities as indicated by robust binding of beads to cells (arrows). Full binding was observed only when epitope I sequences were matching. Homophilic binding of beads to cells was indistinguishable from 'donor' wild-type isoforms, even if epitope II sequences were completely different. In contrast, binding was completely abolished with non-matching epitope I but matching epitope II sequences.

for D109A/N111A/L217A/Q219A. We also mutated several surface-exposed residues of epitope II. In this case, binding assays revealed no difference in adhesion between wild type and the two quadruple mutant isoforms E145A/N147A/M225A/V226A and N147A/F149A/H228A/T229A (Fig. 3a, and Supplementary Fig. 8).

To confirm further that epitope I determines binding specificity, we generated hybrid receptor isoforms swapping either the N-terminal segment of D2 (segment A, containing epitope I; Fig. 3b) or the C-terminal segment of D2 (segment B, containing epitope II; Fig. 3b). We found that hybrid 1/9 (exon 4.1 N terminus; exon 4.9 C terminus) can only bind to isoform Dscam_{1,34,30} but not to isoform Dscam_{9,34,30} (Fig. 3c). In contrast, hybrid 9/1 (exon 4.9 N terminus; exon 4.1 C terminus) will not bind to Dscam_{1,34,30} but binds well to Dscam_{9,34,30}. Hybrid isoforms bearing identical epitope II but different epitope I sequences did not show binding to wild-type isoforms (Fig. 3c). Therefore, homophilic binding specificity in these hybrid receptors is determined by the epitope I sequence.

Because the presented structures document only 2 of the 576 possible exon 4/exon 6 combinations, we sought to potentially discern a general trend by using a phylogenetic comparison, expecting that the proposed functional differences between epitope I and epitope II sequences are reflected in differential sequence conservation. We analysed the sequence conservation²² of epitope I and epitope II among *Drosophila* species and determined the variability among paralogues as well as orthologues (Fig. 4, and Supplementary Figs 11, 12). Alignments of 12 orthologues of the *Drosophila melanogaster* isoform 6.34 show high conservation for residues along the A-A' protrusion of D3 (epitope I sequence), but low conservation of residues along the B strand of D3 (epitope II sequence) (Fig. 4, and

Supplementary Fig. 10). In concurrence, orthologues of isoform 4.1 show high conservation for epitope I along the A' strand of D2, but low conservation for epitope II along the C'-D⁰ unit of D2 (Supplementary Figs 11, 12). The sequences of epitope I segments potentially involved in homophilic interactions are, therefore, highly conserved between species, whereas the sequences that constitute epitope II are more divergent.

In summary, we have provided a structural analysis of the recognition specificity of two variable immunoglobulin domains of *Drosophila* Dscam. Although the D1–D4 structures reported here contain only two variable domains, and it remains to be determined how D7 contributes to binding, our biochemical analysis in the context of the full-length Dscam receptor (Fig. 3) is consistent with an essential contribution of the variable peptide segments of epitope I to the homophilic-binding specificity of Dscam. Swapping the peptide segment containing epitope I but not epitope II resulted in a full switch in binding specificity between two isoforms. This strongly suggests that in a Dscam dimer the matching epitope I peptides enable binding, and non-matching ones inhibit homophilic binding, thereby functioning as a specificity module. The strong sequence conservation of epitope I residues is consistent with a high evolutionary selection pressure preserving a limited set of homophilic-binding interfaces. Although an involvement of epitope II in binding of non-Dscam ligands has not been tested experimentally, the apparently faster-evolving sequence variability in epitope II would be consistent with immune receptor adaptations to dynamic alterations in host–pathogen interactions. We therefore hypothesize that this structural separation of homophilic and heterophilic binding (that is potentially self and non-self recognition) in Dscam may have enabled the parsimonious use of the same gene in creating a large receptor diversity in both the nervous system and immune system.

METHODS SUMMARY

Dscam constructs were expressed both in *Pichia pastoris* and the baculovirus expression system (Invitrogen). For electron microscopy, uranyl formate was used for negative staining, as described¹⁵. The X-ray structure of the D1–D4_{1,34} construct was solved by multiple anomalous dispersion on a K₂PtCl₆ derivative at the X25 beamline of the National Synchrotron Light Source, and refined to a final *R* factor of 17.3% (*R*_{free} = 20.5%). The X-ray structure of the D1–D4_{1,34} construct was solved by molecular replacement and refined to a final *R* factor of 27.1% (*R*_{free} = 30.3%). The sequence analysis is described in the Supplementary Information. The cloning and protein purification of the full-length extracellular mutant constructs of Dscam, as well as the bead aggregation assays, were done as described² with some modification described in detail in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 March; accepted 7 August 2007.

Published online 26 August 2007.

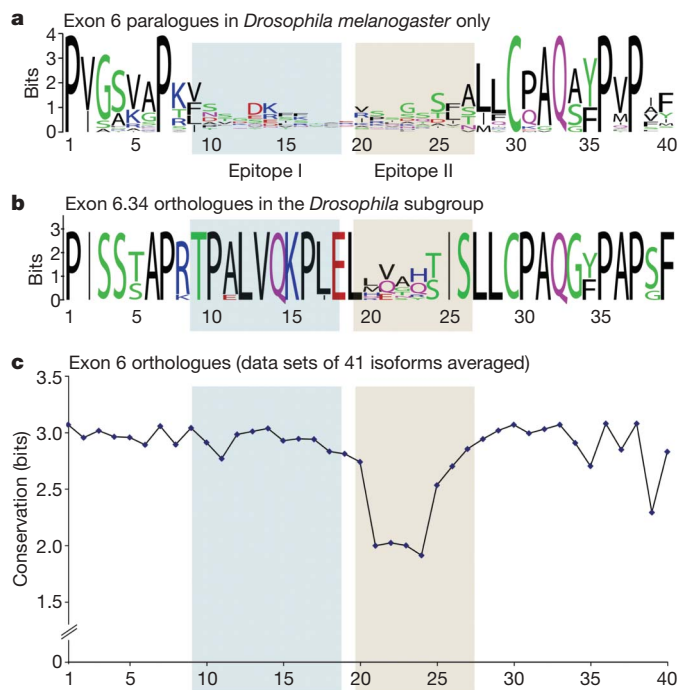


Figure 4 | Phylogenetic comparison reveals differential sequence conservation of epitopes I and II. **a**, Sequence logo representation²² of the conservation of exon 6 isoforms (paralogues) within *Drosophila melanogaster*. A total of 38 isoforms that encode exactly 40 amino acids were used for comparison. Bits, on the y axis, indicates units for evolutionary conservation²². **b**, Sequence logo representation of the conservation of 39 amino acids among orthologues of exon 6.34 in the *Drosophila* subgroup. **c**, Average interspecies conservation of amino acid sites in exon 6. The analysis includes exon 6 isoforms for which orthologies could be established. Average values for 41 data sets (each calculated from at least 8 aligned species) are plotted. Sites with the highest bit score are conserved between all species. Epitopes I and II are shaded in all panels

- Schmucker, D. *et al.* *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* **101**, 671–684 (2000).
- Wojtowicz, W. M., Flanagan, J. J., Millard, S. S., Zipursky, S. L. & Clemens, J. C. Alternative splicing of *Drosophila* Dscam generates axon guidance receptors that exhibit isoform-specific homophilic binding. *Cell* **118**, 619–633 (2004).
- Schmucker, D. & Flanagan, J. G. Generation of recognition diversity in the nervous system. *Neuron* **44**, 219–222 (2004).
- Neves, G., Zucker, J., Daly, M. & Chess, A. Stochastic yet biased expression of multiple Dscam splice variants by individual cells. *Nature Genet.* **36**, 240–246 (2004).
- Watson, F. L. *et al.* Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* **309**, 1874–1878 (2005).
- Wang, J., Zugates, C. T., Liang, I. H., Lee, C. H. & Lee, T. *Drosophila* Dscam is required for divergent segregation of sister branches and suppresses ectopic bifurcation of axons. *Neuron* **33**, 559–571 (2002).
- Hummel, T. *et al.* Axonal targeting of olfactory receptor neurons in *Drosophila* is controlled by Dscam. *Neuron* **37**, 221–231 (2003).
- Zhan, X. L. *et al.* Analysis of Dscam diversity in regulating axon guidance in *Drosophila* mushroom bodies. *Neuron* **43**, 673–686 (2004).
- Zhu, H. *et al.* Dendritic patterning by Dscam and synaptic partner matching in the *Drosophila* antennal lobe. *Nature Neurosci.* **9**, 349–355 (2006).
- Chen, B. E. *et al.* The molecular diversity of Dscam is functionally required for neuronal wiring specificity in *Drosophila*. *Cell* **125**, 607–620 (2006).

11. Hughes, M. *et al.* Homophilic Dscam interactions control complex dendrite morphogenesis. *Neuron* **54**, 417–427 (2007).
12. Matthews, B. *et al.* Dendrite self-avoidance is controlled by Dscam. *Cell* **129**, 593–604 (2007).
13. Soba, P. *et al.* *Drosophila* sensory neurons require Dscam for dendritic self-avoidance and proper dendritic field organization. *Neuron* **54**, 403–416 (2007).
14. Dong, Y., Taylor, H. E. & Dimopoulos, G. AgDscam, a hypervariable immunoglobulin domain-containing receptor of the *Anopheles gambiae* innate immune system. *PLoS Biol.* **4**, e229 (2006).
15. Ohi, M., Li, Y., Cheng, Y. & Walz, T. Negative staining and image classification—powerful tools in modern electron microscopy. *Biol. Proced. Online* **6**, 23–34 (2004).
16. Lo Conte, L., Chothia, C. & Janin, J. The atomic structure of protein–protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198 (1999).
17. Su, X. D. *et al.* Crystal structure of hemolin: a horseshoe shape with implications for homophilic adhesion. *Science* **281**, 991–995 (1998).
18. Freigang, J. *et al.* The crystal structure of the ligand binding module of axonin-1/TAG-1 suggests a zipper mechanism for neural cell adhesion. *Cell* **101**, 425–433 (2000).
19. Schurmann, G., Haspel, J., Grumet, M. & Erickson, H. P. Cell adhesion molecule L1 in folded (horseshoe) and extended conformations. *Mol. Biol. Cell* **12**, 1765–1773 (2001).
20. Harpaz, Y. & Chothia, C. Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J. Mol. Biol.* **238**, 528–539 (1994).
21. Wang, J. Protein recognition by cell surface receptors: physiological receptors versus virus interactions. *Trends Biochem. Sci.* **27**, 122–126 (2002).
22. Schneider, T. D. & Stephens, R. M. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* **18**, 6097–6100 (1990).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Tan for discussion, and E. Reinherz, M. Eck and B. Chen for comments on the manuscript. We also thank R. Zhang and A. Joachimiak at the 19ID beamline of the Advanced Photon Source at the Argonne National Laboratory, and A. Soares at the X25 beamline at the Brookhaven National Synchrotron Light Source for help in X-ray data collection. The molecular electron microscopy facility at Harvard Medical School was established by a generous donation from the Giovanni Armenise Harvard Center for Structural Biology and is maintained by an NIH grant to T.W. We are grateful to S. L. Zipursky and W. Wojtowicz for discussions and the sharing of unpublished results. This work was supported by NIH grants to J.-h.W., D.S. and T.W., and a Pew Scholar Award and John Merck Fund Award to D.S. G.S. is a Damon Runyon fellow, supported by the Damon Runyon Cancer Research Foundation.

Author Information Atomic coordinates and structure factors have been deposited in the Protein Data Bank: 2V5M (Dscam D1–D4_{1,34}, space group *P*4₂2₂), 2V5S (Dscam D1–D4_{1,34}, space group *C*222₁) and 2V5R (Dscam D1–D4_{9,9}, space group *C*2). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.-h.W. (jwang@red.dfci.harvard.edu) and D.S. (dietmar_schmucker@dfci.harvard.edu).

METHODS

Cloning, expression and protein purification for structural studies. Nucleotide sequences encoding Dscam D1–D4_{1,34} and D1–D4₉ were amplified from complementary DNA plasmids by PCR. Two plasmids containing combinations of alternative exons 4 and 6 (4.1&6.34 and 4.9&6.9) were used as templates. An N-terminal His-tag followed by a TEV-protease recognition site and *SacI* and *HindIII* restriction sites were added. Recombinant baculovirus was generated by recombination with Bac-N-Blue DNA by Invitrogen. Virus was amplified in *Spodoptera frugiperda* (SF9) cells (Invitrogen) and High five cells to a density of 1.5×10^6 cells ml⁻¹ and infected with the recombinant baculovirus at a multiplicity of infection between 1 and 10. At 72 h post infection, cells were spun down and the medium was filtered using a 0.22 µm CA membrane (Corning) and dialysed overnight against PBS, pH 7.0, supplemented with 100 mM NaCl. The medium was loaded on a Ni-NTA column (Qiagen) equilibrated with PBS, pH 7.0, 100 mM NaCl and 10 mM imidazole by gravity flow. Recombinant Dscam was eluted with PBS, pH 7.0, 100 mM NaCl and 200 mM imidazole. The His-tag was cleaved with TEV protease (Invitrogen) and subsequently removed by another run over the Ni-NTA column. Any remaining contaminating proteins were removed by gel filtration using a Superdex 200 column. The protein was concentrated to 20 mg ml⁻¹ in 20 mM Hepes buffer, pH 7.4.

The inducible eukaryotic *Pichia pastoris* expression system (EasySelect, Invitrogen) was also used to produce the N-terminal-eight-domain constructs. Dscam D1–D8 constructs were cloned into pPCIZα (Invitrogen) with an N-terminal His(6)-tag followed by a TEV-protease recognition site. For transformation, electrocompetent *Pichia* cells were prepared according to a standard protocol (Invitrogen) and 6–15 µg of the expression plasmid, linearized with *PmeI* (NEB), was added to 80 µl of cells and incubated on ice for 10 min. Cells were electroporated using a Micropulser (Biorad) and cells were resuspended carefully in 1 ml of cold 1 M sorbitol solution. Cells were allowed to recover at 30 °C without shaking for 1 h and another 30 min after addition of 500 µl YPD-medium. Cell suspension (120 µl) was plated on YPD-agar plates containing 100 µg or 300 µg of zeocin, and plates were screened for colonies after 3–4 days.

For expression screening of clones, colonies were picked and resuspended into 4 ml of BMGY-medium in deep 24-well plates (Whatman) and grown for 24 h. Cells were allowed to settle and the medium was exchanged for 1 ml BMMY to induce expression. Medium was harvested and spun after 16–20 h and loaded on a 4–12% Bis-Tris gel under denaturing conditions. For large-scale protein production, strongly expressing clones were grown in 500 ml BMGY for 24 h, spun, and resuspended in 100 ml BMMY. After 20 h, cells were spun down, medium was filtered and the recombinant protein was purified using Ni-NTA (Qiagen) affinity-chromatography. After washing with 12 bed volumes of HEPES, pH 7.5, buffer was supplemented with 100 mM NaCl and 25 mM imidazole. Protein was eluted with HEPES, pH 7.5, 100 mM NaCl and 250 mM imidazole. Typically, yields were 30–60 mg l⁻¹. For removal of the N-terminal His-tag, the protein was incubated for 24 h at room temperature with recombinant TEV-protease (Invitrogen).

Crystallization and structure determination. Crystals of Dscam D1–D4_{1,34} were grown in 1.5 M ammonium sulphate and 0.1 M Hepes buffer at pH 7.5. Three crystal forms were obtained (*P*₄2₂: *a* = *b* = 99.2 Å, *c* = 164.0 Å; *I*₄2₂: *a* = *b* = 146.7 Å, *c* = 325.5 Å; *C*22₁, including the His-tag: *a* = 99.8 Å, *b* = 166.8 Å, *c* = 125.6 Å). Crystals of Dscam D1–D4₉ were grown in 10% PEG 8000, 1 mM spermidine and 0.1 M Tris-HCl at pH 8.5 in space group *C*2 with cell dimensions of *a* = 277.8 Å, *b* = 70.5 Å, *c* = 72.8 Å and β = 105.1°. Native data were collected with a Quantum Q315 CCD detector (ADSC) at the 19ID beamline at the Advanced Photon Source. Multi-wavelength anomalous dispersion data were collected from the *I*₄2₂ crystal form of D1–D4_{1,34} with a Quantum Q315 CCD detector (ADSC) at the X25 beamline at the Brookhaven National Synchrotron Light Source. Three wavelengths were used (see Supplementary Table 1), and data were processed and scaled with HKL2000 (ref. 23). The Pt sites were located with SHELXD²⁴ using data from 12 to 3.7 Å. The phases were refined with SHARP²⁵ to a resolution of 3.5 Å with a Figure of Merit of 0.38. The phased map was solvent-flattened and averaged with DM²⁶ to 3.5 Å resolution along the translation vector (0.0, 0.0, 0.5) as obtained from the native Patterson. Iterative cycles of manual model building with Xtalview²⁷ and multi-crystal averaging with DMMULTI²⁸ were performed and model refinement was switched to the *P*₄2₂ crystal form, extending to 2.0 Å resolution. Arpwrap²⁹ was used to further improve the map by iterative rebuilding (in the molrep mode) and subsequent tracing of the model. A total of 350 out of 388 residues were automatically traced, and 349 side chains docked. The model was completed manually and refined to an *R*-factor of 17.3% (*R*_{free} 20.5%) with Refmac5 (ref. 30) using a bulk solvent correction, including hydrogens and TLS refinement. The final model contains 388 amino acid residues, two glycans and 709 water molecules. The Ramachandran statistics indicate 98.3%

of the residues are in the most-favoured region and 1.7% of the residues are in the generously allowed region. The Dscam D1–D4₉ structure was solved by molecular replacement with MOLREP⁴ using Dscam D1–D4_{1,34} as a search model. The structure was refined at 3.0 Å to an *R*-factor of 27.1% (*R*_{free} 30.3%) with PHENIX³¹ using non-crystallographic symmetry and B group refinement, and contains 391 residues. The Ramachandran statistics indicate 96.8% of the residues are in the most-favoured region and 3.2% of the residues are in the generously allowed region.

Electron microscopy and image processing. Dscam constructs D1–D4 and D1–D8 were prepared for electron microscopy using the conventional negative-staining protocol³². Briefly, 3 µl of sample were adsorbed to a glow-discharged carbon-coated copper grid and stained with 0.75% uranyl formate. The sample was imaged with a Tecnai T12 electron microscope equipped with a LaB₆ filament and operated at an acceleration voltage of 120 kV. Images were recorded using low-dose procedures at a magnification of $\times 52,000$ and a defocus value of about -1.5 µm (Supplementary Fig. 2). Micrographs were digitized with a Zeiss SCAI scanner using a step size of 7 µm. Pixels (2 \times 2 and 3 \times 3) were averaged to obtain a pixel size of 2.7 Å and 4.05 Å on the specimen level for D1–D4 and D1–D8, respectively. For the projection analysis 12,093 D1–D4 particles and 5,799 D1–D8 particles were interactively selected from images using the display program WEB, which is associated with the SPIDER program suite³³. The D1–D4 particles were windowed into 90 \times 90 pixel images and the D1–D8 particles into 120 \times 120 pixel images, and subjected to 10 cycles of multi-reference alignment and K-means classification specifying 50 output classes for D1–D4 and 100 classes for D1–D8.

Sequence analysis. Genomic sequence data were obtained by BLASTN and TBLASTN searches using Flyblast (<http://flybase.net/bblast/>) and the NCBI trace archive. Genomic sequences of the following species were used for the analysis: *D. melanogaster*, *D. simulans*, *D. sechellia*, *D. yakuba*, *D. erecta*, *D. ananassae*, *D. pseudoobscura*, *D. persimilis*, *D. willistoni*, *D. mojavensis*, *D. virilis*, *D. grimshawi*. Protein sequences were aligned by CLUSTALW³⁴ and sequence logos were generated with WebLogo (<http://weblogo.berkeley.edu/>), using a small-sample correction.

Cloning and protein expression of full-length extracellular mutant constructs for bead aggregation assays. The pIB-EC16-Fc and the pcDNA3-Dscam-1.30.30 vectors were a gift from W. Wojtowicz and L. Zipursky. The pIB-EC16-Fc contains a Dscam construct encoding the full-length extracellular domain with the exon combination exon 4.1, exon 6.30 and exon 9.30 (termed 1.34.30) as well as the Fc region of human IgG inserted into the pIB vector (pIB-EC16-Fc). A fragment containing Dscam immunoglobulin-domains d2–d5 was excised from pIB-EC16-Fc by restriction digest with *MfeI* and *AatII*. Dscam isoform 1.34.30 was amplified from a full-length Dscam construct¹⁰, cut with *MfeI* and *AatII* and ligated into pIB-EC16-Fc. All point mutants for d2 and d3 were introduced by overlap PCR using Dscam 1.34.30 as a template (primer sequences are available on request). Dscam exon 6 in the pcDNA3-Dscam-1.30.30 was swapped by restriction digest cloning using *NheI* and *KpnI* sites to obtain pcDNA3-Dscam-1.34.30.

S2 cells in serum-free medium (Gibco) were transfected with pIB-EC16-Fc wild-type and mutant constructs using Cellfectin (Invitrogen). After 3 days, stably transfected cells were selected for by adding 30 µg ml⁻¹ Blasticidine (Sigma). After 1 week of selection, cells were expanded and the medium containing recombinant protein was harvested every two days. EC16-Fc proteins were captured and affinity-purified with protein A Sepharose (Repligen), washed with PBS, and eluted with 100 mM glycine, pH 2.8, followed by neutralization with 1 M Tris, pH 8. Protein integrity and purity was tested by Coomassie staining and protein concentrations were calculated by spectroscopy.

Hybrid isoforms were generated by overlap PCR using the isoforms 1.34.30 and 9.34.30 as template. Vectors, expression and protein purification were identical to wild-type isoforms described above. Primer and plasmid sequences are available on request.

Binding of beads to Cos cells. The binding-assay was carried out essentially as described² with some modifications: Cos-7 cells were grown in DMEM, 10% FCS, plated on round microscope coverslips in 24-well plates, grown to 70–80% confluency, and transfected with 300 ng pcDNA3-Dscam-1.34.30 and 130 ng pcDNA3-GFP per well using the Eugene 6 transfection reagent (Roche).

Red fluorescent sulphate polystyrene microspheres (125 µl; Duke Scientific) with a diameter of 0.39 µm and a stock concentration of 2% solids were washed in PBS and coated with a mouse anti-human IgG Fc antibody (Abazyme) overnight at 4 °C, washed again and resuspended in 250 µl of PBS, 0.1% BSA. Beads were further diluted 20-fold in PBS, 0.1% BSA and briefly sonicated. Twenty microlitres of beads were then incubated in a total volume of 100 µl PBS, 0.1% BSA with 5 µg ml⁻¹ Dscam wild-type or mutant proteins for 2 h at 15 °C with 600 r.p.m. agitation. Dscam-coated beads were sonicated in a cup horn and added to the Cos cells in DMEM containing 10% FCS at a final bead

concentration of 0.002% solids. Cells were allowed to bind beads for 2 h at 37 °C, followed by 5 washing steps with DMEM. Cover slips were mounted on microscope slides for imaging.

Image analysis and quantification. Cells were imaged with a Zeiss Axioskop fluorescence microscope: Random fields with GFP-positive cells were selected and corresponding images of the green (cells) and red (beads) channels were saved. Images were processed and scrambled for blind manual counting. For each field the total number of GFP-expressing cells and GFP-positive, 'binding' cells was determined. A cell was called 'binding' if it was covered with at least ten beads and if the beads resembled the cell shape (that is, a large cell binding beads in only a small area was not called 'binding'). Next, the percentage of all binding cells was calculated by adding total cells and binding cells from all fields for a given mutation. About 20% of cells were binding cells if incubated with 1.34.30 wild-type beads. To aggregate data from experiments that were performed on different days, binding was expressed as 'binding as a per cent of 1.34.30 wild-type binding'. Variation within the experiment was assessed for each mutant protein by calculating the sample standard deviation of the percentage of binding cells for each field. The total number of cells that were counted for each mutation is: K112E, H114D, 1,103; V218P, 1,408; D109A, N111A, L217A, Q219A, 1,986; N111A, K112A, K220A, P221A, 1,895; N111A, K112A, 2,313; L217A, Q219A, 1,760; E145A, N147A, M225A, V226A, 1,620; N147A, F149A, H228A, T229A, 1,755; 1.34.30 wild type, 3,938; 1.30.30 wild type, 241.

23. Otwinowski, Z. & Minor, W. in *Macromolecular Crystallography* (eds Carter, C. W. Jr & Sweet, R. M.) 307–326 (Academic Press, New York, 1997).
24. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1772–1779 (2002).
25. La Fortelle, E. & Bricogne, G. in *Methods in Enzymology, Macromolecular Crystallography* (eds Carter, C. W. Jr & Sweet, R. M.) 472–494 (Academic Press, New York, 1997).
26. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
27. McRee, D. E. XtalView/Xfit—A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165 (1999).
28. Cowtan, K., Zhang, K. & Main, P. in *International Tables for Crystallography* (eds Rossmann, M. & Arnold, E.) 25.2.5 (Kluwer Academic Publishers, Dordrecht, 2001).
29. Perrakis, A., Morris, R. & Lamzin, V. S. Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463 (1999).
30. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
31. Adams, P. D. *et al.* Recent developments in the PHENIX software for automated crystallographic structure determination. *J. Synchrotron Radiat.* **11**, 53–55 (2004).
32. Ohi, M., Li, Y., Cheng, Y. & Walz, T. Negative staining and image classification—powerful tools in modern electron microscopy. *Biol. Proced. Online* **6**, 23–34 (2004).
33. Frank, J. *et al.* SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199 (1996).
34. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).

LETTERS

Crystal structure of the heterotrimer core of *Saccharomyces cerevisiae* AMPK homologue SNF1

Gabriele A. Amodeo^{1*}, Michael J. Rudolph^{1*} & Liang Tong¹

AMP-activated protein kinase (AMPK) is a central regulator of energy homeostasis in mammals and is an attractive target for drug discovery against diabetes, obesity and other diseases^{1–5}. The AMPK homologue in *Saccharomyces cerevisiae*, known as SNF1, is essential for responses to glucose starvation as well as for other cellular processes, although SNF1 seems to be activated by a ligand other than AMP^{1,6–8}. Here we report the crystal structure at 2.6 Å resolution of the heterotrimer core of SNF1. The ligand-binding site in the γ -subunit (Snf4) has clear structural differences from that of the *Schizosaccharomyces pombe* enzyme⁹, although our crystallographic data indicate that AMP can also bind to Snf4. The glycogen-binding domain in the β -subunit (Sip2) interacts with Snf4 in the heterotrimer but should still be able to bind carbohydrates^{10–13}. Our structure is supported by a large body of biochemical and genetic data on this complex^{1,6–8,14–18}. Most significantly, the structure reveals that part of the regulatory sequence in the α -subunit (Snf1)^{15,16,18,19} is sequestered by Snf4, demonstrating a direct interaction between the α - and γ -subunits and indicating that our structure may represent the heterotrimer core of SNF1 in its activated state.

Most AMPKs (including SNF1) are heterotrimeric enzymes, consisting of one catalytic subunit (α) and two regulatory subunits (β and γ ; Fig. 1a, and see Supplementary Figs 1–3 for alignment of their sequences). The Ser/Thr protein kinase domain is located at the amino terminus of the α -subunit^{20,21}, and a regulatory sequence helps to control the activity of this enzyme^{15,16,18,19}. The β -subunit contains a glycogen-binding domain (GBD) that may target the heterotrimer to glycogen storage sites^{10,11,13}. The γ -subunit contains two tandem pairs of the cystathionine- β -synthase (CBS) motifs, each of which is also known as a Bateman domain^{22,23}. Binding of AMP or ATP to the γ -subunit regulates the activity of the protein kinase domain in the α -subunit. This is supported by the observations that mutations in the γ -subunit can affect nucleotide binding and cause diseases^{1,2,24,25}.

We have determined the crystal structure of the heterotrimer core of *S. cerevisiae* SNF1 at 2.6 Å resolution. The expression construct contains residues 398–633 of the α -subunit (Snf1, lacking the protein kinase domain), residues 154–415 of the β -subunit (Sip2), and the entire γ -subunit (Snf4, residues 1–322; Fig. 1a). (We are following the convention of using SNF1 to indicate the heterotrimer and using Snf1 to indicate the α -subunit⁶.) The three subunits were expressed together in *Escherichia coli* from a single plasmid, and they purified together through Ni²⁺-affinity and gel-filtration chromatography (Supplementary Fig. 4). The refined structure has excellent agreement with the crystallographic data and the expected bond lengths, bond angles, and other geometric parameters (Supplementary Table 1). Most of the residues (82%) are in the most favoured region of the Ramachandran plot.

The structure shows that there are intimate interactions among the three subunits in the heterotrimer, as observed previously with the *S. pombe* enzyme⁹. A central component of this heterotrimer interface is an eight-stranded, mostly antiparallel β -sheet, formed with four strands from Snf1 (residues 531–586), three strands from Sip2 (from the extreme C-terminal region, residues 375–412) and one strand from Snf4 (residues 38–45, the first β -strand of CBS1, β 1A; Fig. 1b and Supplementary Fig. 5). The hydrophobic core of this interface is located between the β -sheet and two helices from Snf1 (residues 515–529 and 612–630). The two neighbouring strands of the β -sheets in Snf1 and Sip2 share only a few hydrogen bonds at one end, and are splayed apart from each other at the other end (Fig. 1b and Supplementary Fig. 5). Residues 504–511 of Snf1 form two small β -strands (shown in red in Fig. 1b, because they are part of the regulatory sequence) that interact with the open ends of the two β -sheets, providing further stabilization of the heterotrimer. Additional interactions between the three subunits are mediated by the β -hairpin structures in the β -sheets of the Bateman1 domain of Snf4 (Fig. 1b and Supplementary Fig. 5). Our structure of the heterotrimer core is supported by a large body of biochemical and genetic studies on SNF1 (refs 6, 14, 15, 17).

We have observed new structural features for the AMPK heterotrimer from our studies, because the expression construct for the SNF1 complex contains about 100 more residues for the α -subunit and 160 more residues for the β -subunit than those of the *S. pombe* enzyme structure⁹. Most importantly, the structure shows that residues 460–495 of Snf1 have well-defined interactions with Snf4, including an antiparallel β -sheet between residues 467–469 of Snf1 and residues 270–275 of Snf4 (the first β -strand of CBS4, β 4A; Fig. 1b and Supplementary Fig. 6). This demonstrates a direct interaction between the α - and γ -subunits in SNF1, which is supported by earlier biochemical and genetic data¹⁵. The L470S mutation in Snf1, which can disrupt this interaction¹⁵, is located in the interface between Snf1 and Snf4 in our structure (Fig. 1b and Supplementary Fig. 6). This direct interaction has significant implications for the regulation of the protein kinase activity of SNF1, because residues 460–495 are part of the regulatory sequence of Snf1 (see below).

Another new feature of our structure is the presence of the GBD in the β -subunit (Sip2; Fig. 1a). The GBD has close interactions with the rest of the heterotrimer, primarily with the N-terminal region (the first helix and the following loop) and a helix (α 2B) in CBS2 of Snf4 (Fig. 1b). Residues in the hydrophobic core of this interface are mostly conserved, indicating that the GBD may also have a similar conformation in mammalian AMPKs. The putative binding site for carbohydrates, as revealed by the binding of β -cyclodextrin¹⁰, is exposed to the solvent and should be available for binding ligands in this complex (Fig. 1b).

¹Department of Biological Sciences, Columbia University, New York, New York 10027, USA.

*These authors contributed equally to this work.

Besides these new structural features, the overall structure of the *S. cerevisiae* SNF1 heterotrimer is similar to that of the *S. pombe* AMPK reported recently (Fig. 1c)⁹. The root mean squared distance for 280 equivalent C α atoms in the γ -subunits of the two structures is 1.4 Å, and that for 142 equivalent C α atoms in the α - and β -subunits is 1.2 Å. However, there is a significant difference in the positions of the α - and β -subunits relative to the γ -subunit in the two structures (Fig. 1c), corresponding to a rotation of about 12°. Whereas all the loops in the *S. pombe* structure are defined, many of the loops in the structures of the Snf1 and Sip2 subunits are missing in the current SNF1 structure as a result of disorder (Fig. 1b).

The two Bateman domains in the γ -subunit are arranged in a head-to-head fashion, with an extensive, hydrophobic interface between them formed by two helices (α A and α B) from each CBS motif (Fig. 2a). (The secondary structure elements in the CBS motifs are named in accordance with the convention devised previously²⁶.) The first CBS motif is much larger than the second one in each Bateman domain (Fig. 1a), because it contains an extra crossover connection and an extra helix (α C; Fig. 2a)²⁶. Overall, the γ -subunit has the shape of a disc, with a diameter of about 55 Å and a thickness of about 30 Å.

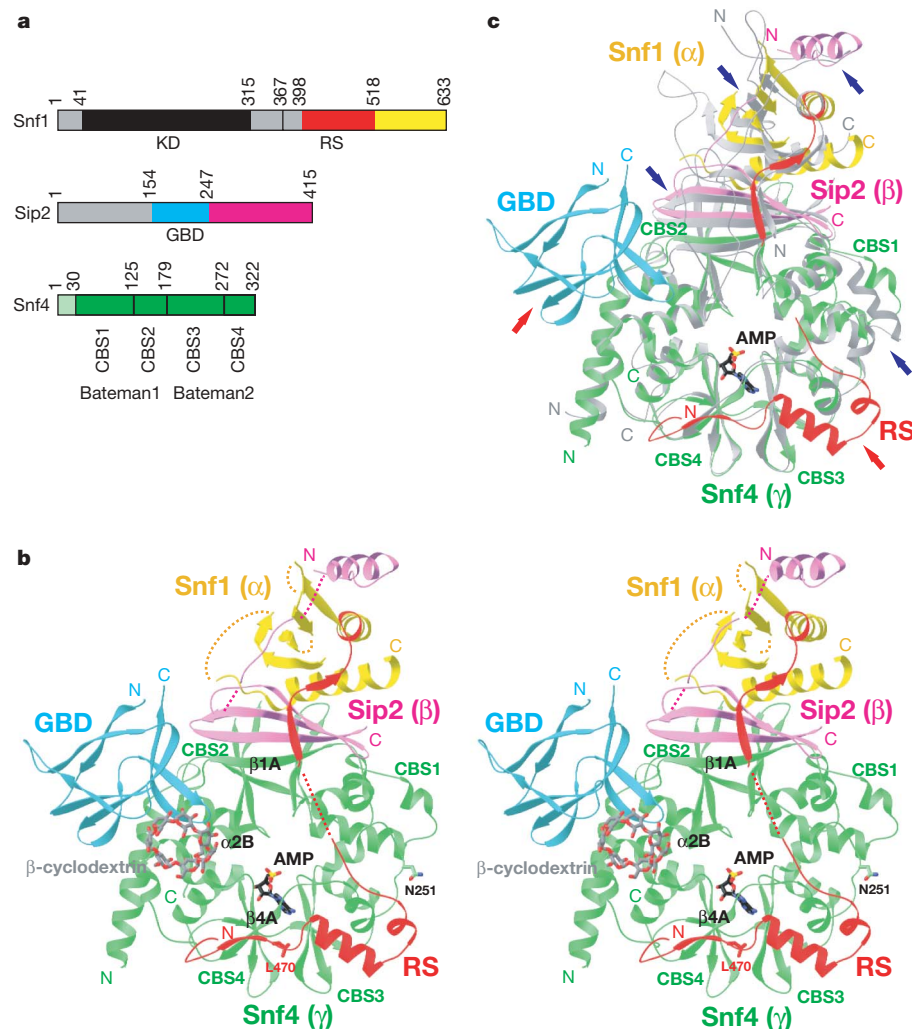
Earlier structural studies of the Bateman2 domain alone, from *S. cerevisiae*²⁶ and from humans²⁷, showed a homodimer, and the overall shape of this dimer is also a disc (Fig. 2b). However, the two monomers are arranged in a head-to-tail fashion, even though the same two helices (α A and α B) are used for the formation of the dimer. The structure of the Bateman2 domain in this homodimer is very different from that in full-length Snf4, with a root mean squared distance of 1.3 Å for 113 equivalent C α atoms (Fig. 2c). The largest

structural differences are for the two helices (α A and α B) that mediate the dimeric association. Nonetheless, the human Bateman2 domain homodimer can bind two AMP molecules, obeying the two-fold symmetry of the dimer²⁷.

We included 1 mM AMP in the crystallization solution but did not observe any electron density for it in this crystal. A comparison of our structure with that of *S. pombe* AMPK in complex with AMP⁹ reveals that there are main-chain and side-chain differences between them in the AMP-binding site (Fig. 3). The phosphate group of AMP interacts with the side chains of three Arg residues in *S. pombe* AMPK⁹ (Fig. 3). One of them is equivalent to Arg 143 in Snf4, whose side chain clashes with the AMP molecule in the *S. pombe* enzyme (Fig. 3). The second Arg residue is equivalent to Gly 145 in Snf4 (Fig. 3) and a His residue in the mammalian γ -subunits (Supplementary Fig. 3). Mutation of this His residue to Gly is sufficient to make the mammalian enzyme insensitive to AMP²⁸. Mutation of this His residue to Arg in the γ 2-subunit is linked to the Wolff–Parkinson–White syndrome and reduces the affinity for the nucleotide²⁴. The third Arg residue, Arg 294 in Snf4, also corresponds to the site of a disease-causing mutation in human AMPK. Two other disease-causing mutations in mammalian γ -subunits, equivalent to Val63 and Thr 166 in Snf4, are located in a pocket on the opposite face of the Snf4 disc (Fig. 3).

Crystallographic analysis on diffraction data at 3.3 Å resolution collected on another crystal, grown under the same conditions but with large differences in unit cell parameters, revealed the presence of an AMP molecule in Snf4 (Supplementary Fig. 7), at essentially the same position as that observed in the *S. pombe* enzyme (Fig. 3). The

Figure 1 | Structure of the heterotrimer core of *S. cerevisiae* SNF1. **a**, Domain organization of SNF1 subunits. Residues that are included in the co-expression construct are shown in colour, and the others are shown in grey or black. KD, protein kinase domain; RS, regulatory sequence. **b**, Schematic representation (stereo view) of the heterotrimer core of SNF1. The regulatory sequence of the α -subunit (Snf1) is shown in red and the rest is in yellow; the GBD of the β -subunit (Sip2) is shown in cyan and the rest is in magenta; and the γ -subunit (Snf4) is shown in green. The positions of AMP (stick model in black), as observed from our studies and in the *S. pombe* enzyme⁹, as well as that of β -cyclodextrin (in grey) as bound in the rat GBD¹⁰, are shown for reference. **c**, Superposition of the structures of *S. cerevisiae* SNF1 (coloured as in **b**) and *S. pombe* AMPK (in grey)⁹. The superposition is based on the γ -subunits only. Red arrows point to new features in the SNF1 structure, and blue arrows point to differences between the two structures. Produced with Ribbons³⁰.



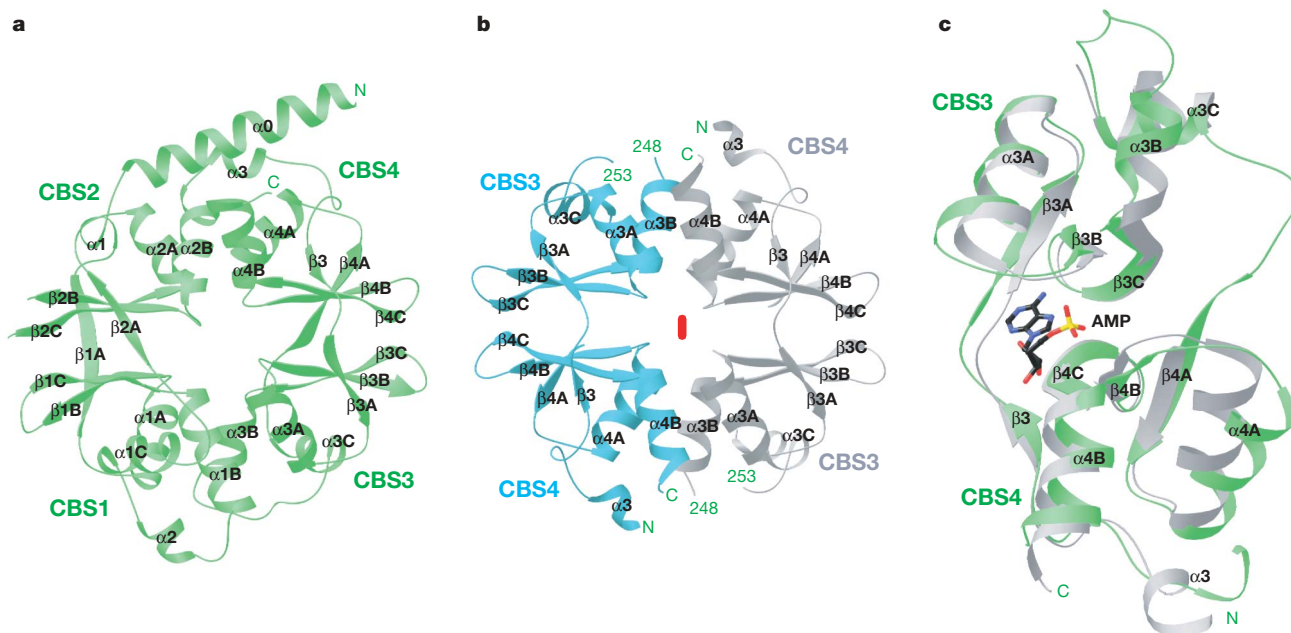


Figure 2 | Large conformational differences for the Bateman2 domain of Snf4. **a**, Structure of the Snf4 subunit, consisting of a Bateman1:Bateman2 'heterodimer'. The secondary structure elements are named in accordance with the system devised earlier²⁶. **b**, Structure of the Bateman2-domain

side chain of Arg 143 is disordered in this complex, and the phosphate group does not seem to have many strong interactions with Snf4. It remains to be determined whether the binding of AMP to Snf4 occurs in solution and why this binding does not seem to affect the activity of the SNF1 heterotrimer.

The *S. pombe* enzyme has been reported to be a dimer in the crystal and in solution⁹. Our crystal of the *S. cerevisiae* enzyme also contains a dimer, although the mode of dimerization is different from that of the *S. pombe* enzyme (Supplementary Fig. 8). The buried surface area for this dimeric association is relatively small, only about 600 Å² per monomer. The heterotrimer is probably monomeric in solution, which is consistent with our gel-filtration data.

dimer of Snf4 (ref. 26). The two monomers are arranged in a head-to-tail fashion. **c**, Overlay of the structures of the Bateman2 domain in full-length Snf4 (in green) and in the homodimer (in grey). Produced with Ribbons³⁰.

Our structure of the SNF1 heterotrimer core has significant implications for the regulation of the protein kinase activity of this enzyme. Previous studies suggested that residues 381–518 of Snf1 represent a regulatory sequence^{15,18}. Residues 313–335 in the $\alpha 1$ -subunit of mammalian AMPK (equivalent to residues 367–392 in Snf1; Supplementary Fig. 1) may auto-inhibit the protein kinase domain, although additional regions also contribute to the regulation^{16,19}. The biochemical data show that the regulatory sequence may control both the inhibition and the activation of SNF1. It interacts with and inhibits the protein kinase domain in the inactive form of AMPK/SNF1, perhaps by binding in the active site groove (Supplementary Fig. 9) or by interacting with the small lobe of the

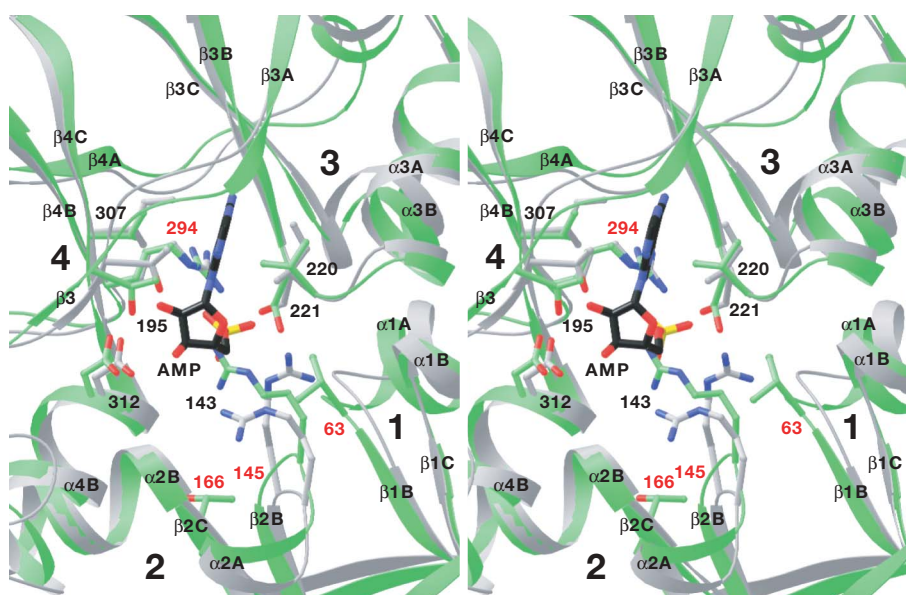


Figure 3 | Structure of the ligand-binding site in *S. cerevisiae* Snf4. Stereo-view overlay of the structures of the γ -subunits of *S. cerevisiae* SNF1 (Snf4, in green) and *S. pombe* AMPK (in grey)⁹. The position of AMP is observed in the *S. pombe* structure⁹, as well as from our studies. Residues that could

interact with AMP are shown, and those that are equivalent to disease-causing mutations in mammalian γ -subunits are labelled in red. Produced with Ribbons³⁰.

protein kinase domain¹⁹. On activation, the regulatory sequence establishes strong interactions with Snf4, thereby relieving the inhibition on the protein kinase domain¹⁵. A simple model based on our structure that is consistent with these data is shown in Supplementary Fig. 9.

Our structure reveals that residues 465–495 in the regulatory sequence are sequestered by Snf4 (Fig. 1b and Supplementary Fig. 6), indicating that this structure might correspond to the heterotrimer core of activated SNF1 (Supplementary Fig. 9). Further support for the importance of this interaction comes from the observation that one of the disease-causing mutations in the γ 2-subunit of mammalian AMPK, namely N488I (equivalent to Asn 251 in Snf4; Supplementary Fig. 3), is located near this part of the regulatory sequence (Fig. 1b). Mammalian AMPKs might therefore use a similar regulatory mechanism, although their sequences in this region are poorly conserved in comparison with Snf1 (Supplementary Fig. 1). Residues 496–518 of the regulatory sequence in Snf1 were found to be required for interaction with the protein kinase domain but not with Snf4 (ref. 15), which is consistent with our structural observations (Fig. 1b). It remains to be determined how the regulatory sequence is recognized by the protein kinase domain in the inactive form of SNF1/AMPK, and how the binding of different ligands to the γ -subunit can control its interaction with the regulatory sequence (Supplementary Fig. 9). Our structure shows that the AMP-binding site and the regulatory sequence are located on opposite faces of the Snf4 disc (Fig. 1b). Our structural information will also be helpful in characterizing, at the molecular level, other mechanisms for the regulation of this heterotrimeric enzyme^{1–4,6,29}.

METHODS SUMMARY

S. cerevisiae Snf1 (residues 398–633), Sip2 (residues 154–415) and Snf4 (residues 1–322) were expressed together by using a tri-cistronic expression system in the pET28a vector (Novagen) (Fig. 1a). The soluble protein was purified by Ni²⁺-affinity and gel-filtration chromatography. The protein was concentrated to 15 mg ml⁻¹ in a solution containing 50 mM Tris-HCl pH 8.5, 150 mM NaCl, 5 mM dithiothreitol and 5% (v/v) glycerol, and then stored at -80 °C.

Crystals of the heterotrimeric complex were obtained by the hanging-drop vapour-diffusion method at 21 °C. The reservoir solution contained 250 mM ammonium citrate pH 7.0 and 15% (w/v) PEG3350. The protein solution was supplemented with 1 mM AMP. The crystals belong to space group C2, with unit cell parameters of $a = 112.3$ Å, $b = 81.8$ Å, $c = 174.7$ Å and $\beta = 102.2^\circ$. There are two copies of the heterotrimer in the crystallographic asymmetric unit. X-ray diffraction data were collected at the X29A beamline of the National Synchrotron Light Source (NSLS). The data processing statistics are summarized in Supplementary Table 1.

The structure was solved by the molecular replacement method, using the structures of the *S. pombe* heterotrimer⁹ as well as the GBD of the rat β 1-subunit¹⁰ as the models. After structure refinement, the models were modified to fit the electron density and sequence of *S. cerevisiae* SNF1, and additional residues that were not in the search models were located. The refinement statistics are summarized in Supplementary Table 1.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 15 May; accepted 27 July 2007.

Published online 12 September 2007.

- Kahn, B. B., Alquier, T., Carling, D. & Hardie, D. G. AMP-activated protein kinase: ancient energy gauge provides clues to modern understanding of metabolism. *Cell Metab.* **1**, 15–25 (2005).
- Hardie, D. G. & Sakamoto, K. AMPK: a key sensor of fuel and energy status in skeletal muscle. *Physiology* **21**, 48–60 (2006).
- Carling, D. AMP-activated protein kinase: balancing the scales. *Biochimie* **87**, 87–91 (2005).
- Kemp, B. E. *et al.* AMP-activated protein kinase, super metabolic regulator. *Biochem. Soc. Trans.* **31**, 162–168 (2003).
- Viollet, B. *et al.* Physiological role of AMP-activated protein kinase (AMPK): insights from knockout mouse models. *Biochem. Soc. Trans.* **31**, 216–219 (2003).
- Hardie, D. G., Carling, D. & Carlson, M. The AMP-activated/SNF1 protein kinase subfamily: metabolic sensors of the eukaryotic cell? *Annu. Rev. Biochem.* **67**, 821–855 (1998).
- Hong, S.-P. & Carlson, M. Regulation of Snf1 protein kinase in response to environmental stress. *J. Biol. Chem.* **282**, 16838–16845 (2007).
- Sanz, P. Snf1 protein kinase: a key player in the response to cellular stress in yeast. *Biochem. Soc. Trans.* **31**, 178–181 (2003).
- Townley, R. & Shapiro, L. Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. *Science* **315**, 1726–1729 (2007).
- Polekhina, G. *et al.* Structural basis for glycogen recognition by AMP-activated protein kinase. *Structure* **13**, 1453–1462 (2005).
- Polekhina, G. *et al.* AMPK β subunit targets metabolic stress sensing to glycogen. *Curr. Biol.* **13**, 867–871 (2003).
- Wiatrowski, H. A. *et al.* Mutations in the Gal83 glycogen-binding domain activate the Snf1/Gal83 kinase pathway by a glycogen-independent mechanism. *Mol. Cell Biol.* **24**, 352–361 (2004).
- Hudson, E. R. *et al.* A novel domain in AMP-activated protein kinase causes glycogen storage bodies similar to those seen in hereditary cardiac arrhythmias. *Curr. Biol.* **13**, 861–866 (2003).
- Iseli, T. J. *et al.* AMP-activated protein kinase β subunit tethers α and γ subunits via its C-terminal sequence (186–270). *J. Biol. Chem.* **280**, 13395–13400 (2005).
- Jiang, R. & Carlson, M. Glucose regulates protein interactions within the yeast SNF1 protein kinase complex. *Genes Dev.* **10**, 3105–3115 (1996).
- Crute, B. E., Seefeld, K., Gamble, J., Kemp, B. E. & Witters, L. A. Functional domains of the α 1 catalytic subunit of the AMP-activated protein kinase. *J. Biol. Chem.* **273**, 35347–35354 (1998).
- Jiang, R. & Carlson, M. The Snf1 protein kinase and its activating subunit, Snf4, interact with distinct domains of the Sip1/Sip2/Gal83 component in the kinase complex. *Mol. Cell Biol.* **17**, 2099–2106 (1997).
- Leech, A., Nath, N., McCartney, R. R. & Schmidt, M. C. Isolation of mutations in the catalytic domain of the Snf1 kinase that render its activity independent of the Snf4 subunit. *Eukaryot. Cell* **2**, 265–273 (2003).
- Pang, T. *et al.* Conserved α -helix acts as autoinhibitory sequence in AMP-activated protein kinase α subunits. *J. Biol. Chem.* **282**, 495–506 (2007).
- Rudolph, M. J., Amodeo, G. A., Bai, Y. & Tong, L. Crystal structure of the protein kinase domain of yeast AMP-activated protein kinase Snf1. *Biochem. Biophys. Res. Commun.* **337**, 1224–1228 (2005).
- Nayak, V. *et al.* Structure and dimerization of the kinase domain from yeast Snf1, a member of the Snf1/AMPK protein family. *Structure* **14**, 477–485 (2006).
- Bateman, A. The structure of a domain common to archaeobacteria and the homocystinuria disease protein. *Trends Biochem. Sci.* **22**, 12–13 (1997).
- Kemp, B. E. Bateman domains and adenosine derivatives form a binding contract. *J. Clin. Invest.* **113**, 182–184 (2004).
- Scott, J. W. *et al.* CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. *J. Clin. Invest.* **113**, 274–284 (2004).
- Hamilton, S. R. *et al.* An activating mutation in the γ 1 subunit of the AMP-activated protein kinase. *FEBS Lett.* **500**, 163–168 (2001).
- Rudolph, M. J. *et al.* Structure of the Bateman2 domain of yeast Snf4: dimeric association and relevance for AMP binding. *Structure* **15**, 65–74 (2007).
- Day, P. *et al.* Structure of a CBS-domain pair from the regulatory γ 1 subunit of human AMPK in complex with AMP and ZMP. *Acta Crystallogr. D* **63**, 587–596 (2007).
- Adams, J. *et al.* Intrasteric control of AMPK via the γ 1 subunit AMP allosteric regulatory site. *Protein Sci.* **13**, 155–165 (2004).
- Sanders, M. J., Grondin, P. O., Hegarty, B. D., Snowden, M. A. & Carling, D. Investigating the mechanism for AMP activation of the AMP-activated protein kinase cascade. *Biochem. J.* **403**, 139–148 (2007).
- Carson, M. Ribbon models of macromolecules. *J. Mol. Graph.* **5**, 103–106 (1987).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Carlson for helpful discussions; H. Robinson and N. Whalen for setting up the X29A beamline; and J. Schwanof and R. Abramowitz for setting up the X4C beamline at the National Synchrotron Light Source. This research is supported in part by an NIH grant to L.T. G.A.A. was supported by an NIH training program in molecular biophysics.

Author Information The atomic coordinates are deposited at the Protein Data Bank under accession number 2QLV. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.T. (ltong@columbia.edu).

METHODS

Protein expression and purification. The three subunits of *S. cerevisiae* SNF1 (residues 398–633 of Snf1, residues 154–415 of Sip2 and residues 1–322 of Snf4) were expressed together using a polycistronic expression system in the pET28a vector (Novagen) (Fig. 1a). The boundaries of each subunit were defined by sequence and secondary structure analysis as well as by experimental sampling of different start sites for the Snf1 and Sip2 subunits. Sip2 carried an N-terminal hexahistidine tag. The proteins were overexpressed in *E. coli* BL21 (DE3) Rosetta cells at 20 °C and purified by Ni²⁺-affinity and gel-filtration chromatography. The protein was concentrated to 15 mg ml⁻¹ in a buffer containing 50 mM Tris-HCl pH 8.5, 150 mM NaCl, 5 mM dithiothreitol and 5% (v/v) glycerol. The protein sample was divided into small aliquots, flash-frozen in liquid nitrogen and then stored at -80 °C. The N-terminal hexahistidine tag was not removed for crystallization. SDS gel electrophoresis on the purified protein showed that the three subunits are present at roughly equal molar concentrations (Supplementary Fig. 1b).

The selenomethionyl protein was produced by using a methionine auxotroph in the defined LeMaster medium³¹, and purified by following the same protocol as that for the native protein.

Protein crystallization. Crystals of the *S. cerevisiae* SNF1 heterotrimer core were obtained by the hanging-drop vapour diffusion method at 21 °C. The reservoir solution contained 250 mM ammonium citrate pH 7.0 and 15% (w/v) PEG3350, and 1 mM AMP was included in the protein solution. Micro-seeding was crucial for producing crystals of sufficient size for structural studies. The crystals were cryoprotected in the reservoir solution supplemented with 12.5% (w/v) PEG3350, and flash-frozen in liquid nitrogen for data collection at 100 K. The crystals belong to space group C2, and the unit cell parameters of the crystal used for data collection are $a = 112.3 \text{ \AA}$, $b = 81.8 \text{ \AA}$, $c = 174.7 \text{ \AA}$, and $\beta = 102.2^\circ$. There are two copies of the SNF1 heterotrimer core in the crystallographic asymmetric unit.

Data collection and processing. A native data set to 2.6 Å resolution and a selenomethionyl single-wavelength anomalous diffraction (SAD) data set to 3.5 Å resolution were collected on an ADSC CCD at the X29A beamline of the National Synchrotron Light Source (NSLS). The diffraction images were processed and scaled with the HKL package³². The data processing statistics are given in Supplementary Table 1.

Structure determination and refinement. The structure of SNF1 heterotrimer core was solved by the molecular replacement method with the program Phaser³³. The structures of the *S. pombe* heterotrimer³⁴ as well as the GBD of the rat $\beta 1$ -subunit³⁵ were used as the models. The structure refinement was performed with the programs CNS³⁶ and Refmac³⁷. The models were manually modified to fit the electron density as well as the sequence of *S. cerevisiae* SNF1,

using the programs O³⁸ and Coot³⁹. Additional residues that were not in the search models were also located. Non-crystallographic symmetry restraints were applied at the beginning stages of the refinement, but were released in the later stages with only minor effects on the *R* and free *R* values. The correctness of the structure solution was confirmed by anomalous difference electron density maps calculated based on the selenomethionyl SAD data. The refinement statistics are presented in Supplementary Table 1.

Observation of AMP binding to the SNF1 heterotrimer. Crystals were obtained by the hanging-drop vapour diffusion method at 21 °C. The reservoir solution contained 210 mM ammonium citrate pH 7.0 and 20% (w/v) PEG3350. The protein solution was supplemented with 1 mM AMP. Micro-seeding was crucial for producing crystals of sufficient size for structural studies. The crystals were cryoprotected in the reservoir solution supplemented with 20% (w/v) PEG3350 as well as 10 mM glucose 6-phosphate for about 16 h, and flash-frozen in liquid nitrogen for data collection at 100 K. The crystals belong to space group C2, with unit cell parameters of $a = 124.2 \text{ \AA}$, $b = 74.7 \text{ \AA}$, $c = 195.4 \text{ \AA}$, and $\beta = 107.4^\circ$. There are two copies of the heterotrimer in the crystallographic asymmetric unit. An X-ray diffraction data set to 3.3 Å resolution was collected at the X4C beamline of the NSLS. The structure was solved by molecular replacement. After crystallographic refinement in the absence of any ligand, the $2F_o - F_c$ electron density seems consistent with the presence of AMP.

31. Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9**, 1665–1672 (1990).
32. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
33. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D* **60**, 432–438 (2004).
34. Townley, R. & Shapiro, L. Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. *Science* **315**, 1726–1729 (2007).
35. Polekhina, G. et al. Structural basis for glycogen recognition by AMP-activated protein kinase. *Structure* **13**, 1453–1462 (2005).
36. Brunger, A. T. et al. Crystallography & NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
37. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
38. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
39. Emsley, P. & Cowtan, K. D. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

LETTERS

Structural basis for AMP binding to mammalian AMP-activated protein kinase

Bing Xiao¹, Richard Heath^{1,2}, Peter Saiu¹, Fiona C. Leiper², Philippe Leone¹, Chun Jing¹, Philip A. Walker¹, Lesley Haire¹, John F. Eccleston¹, Colin T. Davis¹, Stephen R. Martin¹, David Carling² & Steven J. Gamblin¹

AMP-activated protein kinase (AMPK) regulates cellular metabolism in response to the availability of energy and is therefore a target for type II diabetes treatment¹. It senses changes in the ratio of AMP/ATP by binding both species in a competitive manner². Thus, increases in the concentration of AMP activate AMPK resulting in the phosphorylation and differential regulation of a series of downstream targets that control anabolic and catabolic pathways^{1,2}. We report here the crystal structure of the regulatory fragment of mammalian AMPK in complexes with AMP and ATP. The phosphate groups of AMP/ATP lie in a groove on the surface of the γ domain, which is lined with basic residues, many of which are associated with disease-causing mutations. Structural and solution studies reveal that two sites on the γ domain bind either AMP or Mg·ATP, whereas a third site contains a tightly bound AMP that does not exchange. Our binding studies indicate that under physiological conditions AMPK mainly exists in its inactive form in complex with Mg·ATP, which is much more abundant than AMP. Our modelling studies suggest how changes in the concentration of AMP ([AMP]) enhance AMPK activity levels. The structure also suggests a mechanism for propagating AMP/ATP signalling whereby a phosphorylated residue from the α and/or β subunits binds to the γ subunit in the presence of AMP but not when ATP is bound.

AMPK has been described as a cellular fuel gauge³. As early as 1980, the enzyme that was subsequently named AMPK was shown to be regulated by AMP⁴. It is now becoming clear that AMPK has a central role in controlling whole-body metabolism in response to nutrients and hormonal signals¹. AMPK is involved in the regulation of lipid metabolism⁵, feeding and body weight⁶, glucose homeostasis⁷ and mitochondrial biogenesis⁸. In addition to metabolic stress, AMPK is activated by hormones, including leptin⁹, adiponectin¹⁰ and ciliary neurotrophic factor¹¹. AMPK is also activated by metformin¹², one of the most widely used drugs for the treatment of type II diabetes, prescribed to over 120 million people worldwide. Moreover, a recent study suggests that AMPK activation is essential for the therapeutic effects of this drug in the liver¹³. In mammalian cells, AMPK is activated by increases in intracellular AMP by an allosteric mechanism and by regulating the level of AMPK phosphorylation by inhibiting the dephosphorylation of Thr 172 in the activation loop of the kinase domain¹⁴. AMPK is a heterotrimeric complex, consisting of an α catalytic subunit (two isoforms) with β (two isoforms) and γ (three isoforms) regulatory subunits¹⁵. The β subunit acts as a scaffold binding both the α and γ subunits¹⁶, and the carboxy-terminal 85 residues of β is sufficient to form a functional AMP-dependent heterotrimer^{17,18}. Recently, the structure of the AMPK homologue from *Schizosaccharomyces pombe* has revealed the architecture of the heterotrimer subunit interactions¹⁹ but seems to be fundamentally

different from the mammalian enzyme in the way that it senses AMP/ATP.

Using a tri-cistronic expression system we were able to express and purify full-length mammalian AMPK. Because extensive trials failed to yield crystals of full-length material, we produced a range of truncated proteins designed from results of limited proteolysis and sequence analysis. We measured diffraction data extending to 2.1 Å Bragg spacing from crystals obtained from the trimeric complex of the C-terminal domains of α 1(396–548) and β 2(187–272) with full-length γ 1 (Fig. 1a). The structure was solved by molecular replacement and relevant crystallographic statistics are given in Supplementary Table 1.

As the two views of the overall structure of the enzyme in Fig. 1b show, it can be regarded as being made up of two components: an $\alpha + \beta$ module at the bottom interacting with one shoulder of the γ domain module at the top. The interface between the $\alpha + \beta$ and γ modules is mediated by the last two strands of the β subunit forming an interdomain sheet with a β -strand from γ (secondary structure elements are shown on a sequence alignment in Supplementary Fig. 1). The γ subunit, made up of four cystathionine β -synthase (CBS) motifs, is shaped like a flattened disk about 60 Å across and 30 Å deep. These CBS motifs, packing together to generate two Bateman domains, obey approximate 2,2,2 point group symmetry. Each Bateman domain (CBS1 + 2 and CBS3 + 4) arises from the packing of a pair of anti-parallel strands, one from each CBS domain, according to a local dyad axis (Fig. 1b). It is this interaction that generates the adenyl binding sites (Fig. 1c and Supplementary Fig. 2). The two Bateman domains pack around a second dyad axis through an interface dominated by a pair of α -helices from each CBS domain. Passing through the centre of the γ domain there is a small solvent-accessible channel.

Comparison of the structure of mammalian AMPK with its homologue from yeast reveals three striking differences. First, the mammalian enzyme, both in the crystal and in solution, is a monomer (Supplementary Fig. 3), whereas the yeast enzyme has been reported to be a dimer^{19,20}. Second, there is hinging movement between the $\alpha + \beta$ and γ modules such that there is about a 12° difference in their relative orientations between the two species (Supplementary Fig. 4). Third, mammalian AMPK binds three AMP molecules; one does not exchange but the other two compete for binding with Mg·ATP and are responsible for the adenyl-sensing properties of the mammalian enzyme. In the case of the yeast homologue the structure shows one adenyl-binding site, equivalent to the non-exchangeable site in our structure, which seems to bind either AMP or non Mg²⁺-bound ATP¹⁹.

The symmetry relating the CBS domains of mammalian γ means that there are four potential adenyl-binding sites—two at

¹MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK. ²MRC Clinical Sciences Centre, Hammersmith Hospital Campus, Imperial College, DuCane Road, London W12 0NN, UK.

the interface of the two CBS motifs making up each Bateman domain. There is well-defined electron density for AMP in three of these sites but not the fourth. Figure 1c shows some of the common features of these AMP-binding sites. The nucleotide binds in a surface pocket at the interface of the two CBS domains, within each Bateman domain. The adenine moiety sits in a hydrophobic pocket making hydrogen bonds with main-chain groups from two different strands. The phosphate group interacts with the basic side chains of a number of different residues (including Arg 69, His 150, Arg 151, Lys 169, His 297 and Arg 298; see later) and with the hydroxyl groups of either serine or threonine residues. Importantly, there are some interactions of the bound AMP molecules with protein residues from the alternate Bateman to which the binding site is predominantly generated (see later; for example, Lys 169 and His 150). The 2' and 3' hydroxyl groups of the ribose group make a bi-dentate interaction with an aspartic acid residue located on the first turn of an α -helix adjacent to the site (Asp 244 with AMP-1, Asp 89 with AMP-2, and Asp 316 with AMP-3). In the fourth potential adenyl-binding site there is an arginine residue, at position 170, instead of this aspartic acid, which probably accounts for the fact that AMP does not bind at this fourth site.

To understand better the regulation of AMPK by changes in AMP/ATP ratio we carried out a series of binding studies. From the titration studies described in the next paragraph we observed two AMP binding sites on AMPK, apparently in contradiction to the three molecules seen in our crystal structure. We therefore analysed our purified, overexpressed protein by high-performance

liquid-chromatography for bound nucleotides (Methods and Supplementary Fig. 5) and discovered that for both the crystallization fragment and full-length protein there was approximately one adenyl molecule present for every molecule of AMPK. Most of the adenyl species (85%) was present as AMP with a small amount of ADP (15%). As yet we have not managed to displace this co-purifying adenyl moiety and are therefore unable to determine its binding affinity. However, the fact that it persists through the various stages of protein purification, including gel filtration, suggests that the binding is very tight.

We synthesized *N*-methylantraniloyl (mant)-labelled analogues of AMP and ATP to use as probes for determining the affinity of the unlabelled nucleotides. AMPK binds two mant-nucleotides with dissociation constants of $4 \pm 0.5 \mu\text{M}$ (mean \pm s.d.; $n = 3$) (mant-ATP) and $10 \pm 2 \mu\text{M}$ (mant-AMP) (see Supplementary Fig. 6). Dissociation constants for the unlabelled nucleotides were determined using competition assays in which a mant-nucleotide plus unlabelled nucleotide (either AMP or ATP) was titrated with AMPK (see Fig. 2 and Supplementary Fig. 6). The values determined were $12.5 \pm 4 \mu\text{M}$ for ATP and $24.5 \pm 4.5 \mu\text{M}$ for AMP. We also obtained very similar results from titrations carried out with full-length AMPK (Supplementary Fig. 6). Given the recent description of the structure of the AMPK homologue from *S. pombe* in complex with Mg^{2+} -free ATP we investigated the effect of the metal ion on nucleotide binding with mammalian AMPK. Initial experiments (Supplementary Fig. 6) indicated that Mg^{2+} reduced the apparent affinity of AMPK for ATP, however, this effect is only observed at

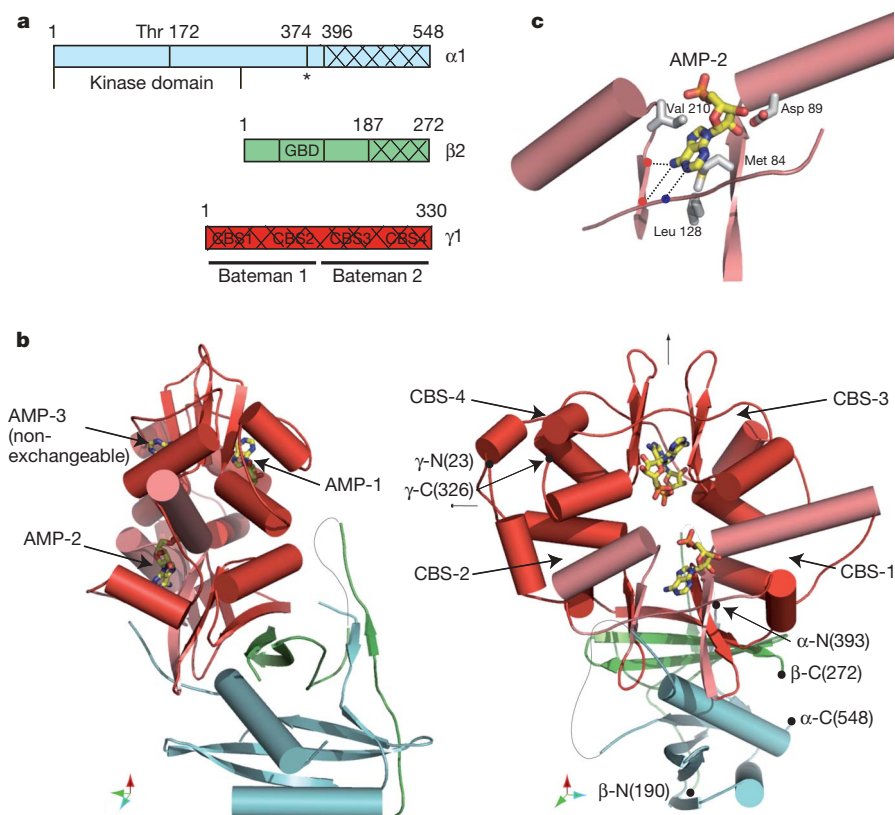


Figure 1 | Structure of mammalian AMPK. **a**, Diagram representation of the three components of the heterotrimer; the constructs used for crystallization are shown hatched. **b**, Ribbon representation of the crystallized complex with three bound AMPs in two orthogonal views. The three domains are coloured according to **a** and the AMP molecules are shown in ball-and-stick representation. For ease of viewing the ends of two loops that are disordered in the structure (α 470–523 and β 223–232) are joined-up by a thin black line, also the first 12 residues of β that interact with a neighbouring molecule in

the crystal lattice are omitted. On the right-hand panel, thin black arrows indicate the position of the approximate two-fold axes relating the four CBS motifs. The γ subunit comprises Bateman domain 1 (CBS1 + 2) and Bateman domain 2 (CBS3 + 4). The interface between the two CBS motifs of each Bateman domain generates two potential adenyl-binding sites, each being made up from a strand and a helix from each CBS motif—these elements, for the AMP-2 site, are coloured in a lighter shade in **b** and are shown in more detail in **c**.

very high concentrations of the metal ion, which are not physiologically relevant. As Fig. 2 shows, titrations of ATP carried out in the presence of 125 μM Mg^{2+} (which will saturate most of the ATP) are the same as those done in the absence of the metal ion. Therefore, at least for mammalian AMPK, the enzyme binds ATP or $\text{Mg}\cdot\text{ATP}$ with the same affinity. This means that under physiological conditions (where nearly all ATP is complexed with Mg^{2+} and there is about 0.4 mM free Mg^{2+} ; ref. 21) the majority of the enzyme will be in complex with two $\text{Mg}\cdot\text{ATP}$ moieties. Mammalian AMPK is therefore different to proteins, such as kinases and small GTPases, whose affinity for the nucleotide is strongly dependent on Mg^{2+} binding, but it is also different to its homologue from *S. pombe* where it is inferred that ATP-bound metal counter ions must be stripped before binding to AMPK¹⁹. Our results are consistent with earlier studies of adenylylation to AMPK γ subunit²², which concluded that the affinity of the protein for ATP was independent of Mg^{2+} .

Only small crystals were obtained from co-crystallization of AMPK with ATP so we soaked ATP into our AMPK/AMP₃ crystals. To test the results of our solution studies, we carried out soaking experiments with $\text{Mg}\cdot\text{ATP}$ and with Mg^{2+} -free ATP. In both cases, the resulting electron-density maps revealed that ATP replaced two out of the three molecules of AMP and that the metal-bound and metal-free forms of the ATP adopted very similar conformations. Consistent with the results of our binding studies, which showed no difference in affinity between the two forms, the magnesium ions complexed between the β and γ phosphate oxygens of the ATP do not interact with the protein. Two views of the structure of the γ subunit with bound nucleotide from both the AMPK/AMP₃ and AMPK/AMP₁(MgATP)₂ complexes are shown as a surface representation in Fig. 3. The two molecules of ATP bind in a similar manner to the AMP molecules that they replace. The β - and γ -phosphates adopt an extended conformation within the binding channel, necessitating the rearrangement of the side-chains of a few basic residues; otherwise there are no significant differences in the two different complexes.

Comparison of the AMP and ATP structures suggests that AMP-3, which does not exchange for ATP, represents the tightly bound adenyly moiety that co-purifies with the protein and, as noted above,

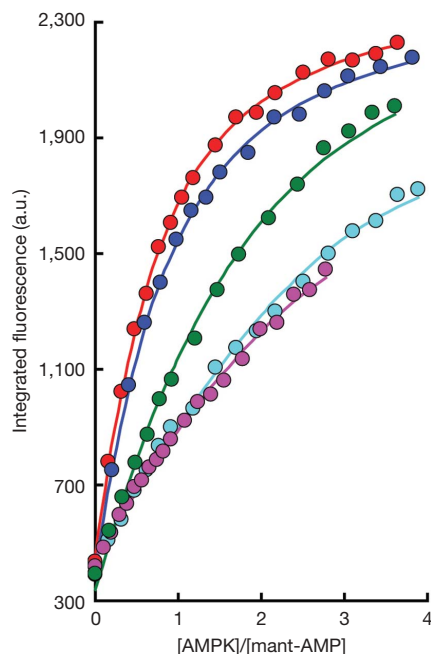


Figure 2 | Fluorescence titrations of mant-AMP with AMPK. Mant-AMP (10 μM) with AMPK; no additions (red), 1.25 mM MgCl_2 (blue), 50 μM AMP (green), 50 μM ATP (magenta), or 50 μM ATP + 125 μM MgCl_2 (cyan).

AMP-3 occupies the equivalent site to the single molecule of AMP seen bound to the yeast homologue¹⁹. All three adenyly-binding sites are structurally related, but comparison suggests that the positioning of two serine residues (225 and 315) that interact with the phosphate group, and a threonine residue (199) that interacts with the ribose, which does not occur in the other two sites, may account for the enhanced affinity of this site. Interestingly, the only one of these residues not conserved in the *S. pombe* sequence is Ser 315, which is an alanine residue in that sequence (C307 in Protein Data Bank 2OOX). At least for mammalian AMPK, our results show that AMP binding at this non-exchangeable site is not involved in AMP/ATP sensing. Clearly mutagenic and biophysical/biochemical studies will be required to elucidate the role of this tightly bound adenyly species.

Taken together, our studies indicate that under physiological conditions the majority of AMPK will contain $\text{Mg}\cdot\text{ATP}$ and only a small proportion of the enzyme will have AMP bound at the two regulatory binding sites. How then do changes in the cellular ATP/AMP ratio affect AMPK activity? To address this issue we set up a simple model to simulate the enzyme's activity, using the binding constants determined here together with data from the literature concerning AMPK activation and adenyly concentrations (Supplementary Fig. 7). Key assumptions of the model are that AMP binding to the enzyme causes allosteric activation as well as inhibiting dephosphorylation of the enzyme's T-loop, and that the non-phosphorylated form has negligible activity. The model shows that 2–3-fold increases in the low micromolar concentrations of AMP bring about a similar fold enhancement in the proportion of AMP-bound enzyme in the presence of a much higher, and essentially unchanging, concentration of

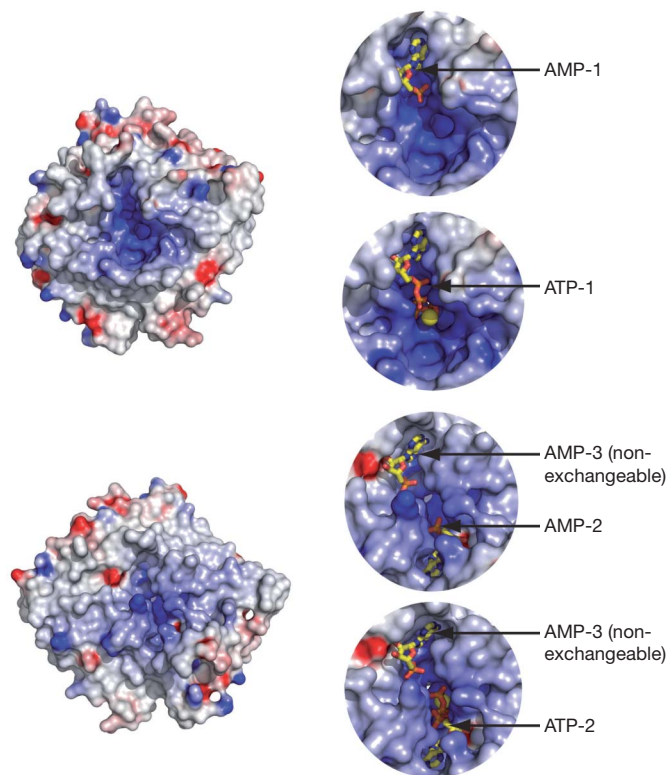


Figure 3 | Front and back views of the three adenyly moieties bound to the γ domain, shown in surface representation. The surface is coloured according to electrostatic potential (blue, positive; red, negative) and the labelled nucleotides are in ball-and-stick representation, with the magnesium ions shown as yellow spheres. The lower panel represents approximately the same orientation of the molecule shown in the right-hand panel of Fig. 1b, whereas the upper panel is viewed from the back. The left-hand panels show the γ domain from the AMP complex, whereas the right-hand panels focus on the phosphate-binding groove from the AMP₃ and ($\text{Mg}\cdot\text{ATP}$)₂AMP₁ complexes.

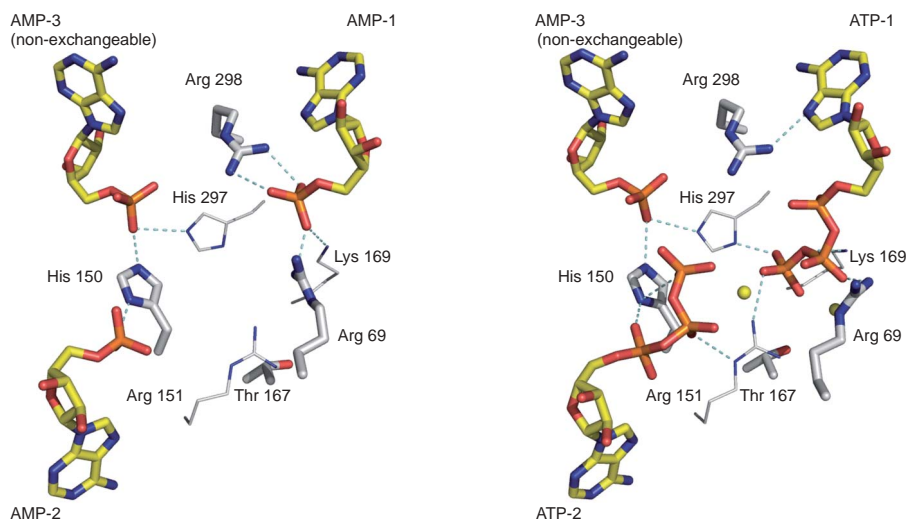


Figure 4 | Interactions between primarily basic residues from γ with the phosphate groups of the three bound adenyl species. The orientation of the molecule is approximately equivalent to that shown in the right-hand panel of Fig. 1b. The left-hand panel is from the AMP₃ complex, and the right-hand panel from the (Mg·ATP)₂AMP₁ complex; the magnesium ions are shown as yellow spheres. In different data sets of the AMP₃ complex, Arg 69 is observed to adopt different side-chain conformations. The highest resolution structure is shown here, but Arg 69 adopts a conformation that

results in it interacting with AMP-2 in another data set and we take this as evidence for a low energy difference for these two conformations. Arg 151 interacts with both ATP-1 and -2 but not with the AMP counterparts. Residues whose mutation is implicated in glycogen storage cardiomyopathy are shown with thicker bonds. The equivalent residues in $\gamma 2$ are: Arg 69=Arg 302, His 150=His 383, Thr 167=Thr 400 and Arg 298=Arg 531 (see Supplementary Table 2).

Mg·ATP. Importantly, this change in AMP-bound AMPK accounts for the 2–3-fold enhancement of muscle AMPK activity observed following moderate exercise^{23,24}. A key prediction of this model is that only a small proportion of AMPK is active at any one time. This notion, of a large pool of essentially inactive AMPK, is supported by studies using prolonged exercise²⁵ or non-physiologic methods to activate AMPK; for example, the use of respiratory uncouplers that generate a large increase in AMP and lead to increases in enzyme activity at least 5–10-fold greater than that stimulated by moderate exercise²⁶.

A glycogen storage cardiomyopathy associated with an electrical conductance disorder in humans^{27–29} has been attributed to naturally occurring mutations in the $\gamma 2$ isoform of AMPK. These mutations generally interfere with the normal activation of AMPK by AMP^{14,22,30}. Mapping of the corresponding residues onto the $\gamma 1$ structure reveals that eight out of the ten mutations that have been characterized involve amino acids whose side chains are in proximity to the adenyl-binding sites (Fig. 4a, b and Supplementary Table 2). Six of these are seen to be contacting the phosphate groups of the AMP/ATP. The structure also shows how small changes in the side-chain conformations of these residues accommodate either AMP or ATP in the binding site. All the mutations tested for adenyl binding (Supplementary Table 2) have been shown to reduce the enzyme's affinity for both AMP and ATP, in some cases quite substantially²². All but one of these mutations is identified here as contacting the adenyl phosphate groups and the remaining one (a leucine insertion) is close to this site and may well affect binding also. Once the molecular basis for signal propagation by AMP binding to AMPK is further elucidated, additional insights into the actions of these mutations may be revealed.

As discussed in our modelling study, once phosphorylated, AMPK has substantial basal activity that can be modestly enhanced by the allosteric effect of AMP binding. However, AMP binding also significantly decreases the rate of dephosphorylation of activated AMPK. Any molecular model of AMPK action must account for both of these phenomena. We think the lack of any significant conformational change in the γ subunit between the AMP and ATP bound forms implies that the activation signal, initiated by AMP binding, is not propagated by the limited rearrangements of some basic side chains

at the binding site. Instead, we hypothesize that AMP binding triggers the formation of inter-subunit interactions that are not possible when ATP is bound. Such changes in the quaternary structure of the heterotrimer then influence the activity of the kinase domain and its susceptibility to dephosphorylation. Given the structure and charge associated with the ATP-binding sites, a plausible binding partner for AMP would be a phosphorylated residue from either the α and/or β subunits (at least eight such phosphorylation sites have been identified). The fact that our structure shows there are two exchangeable AMP/ATP sites, on opposite faces of the γ subunit, suggests the possibility that two distinct inter-subunit interactions with γ occur, perhaps enabling separate mechanisms for the allosteric activation and dephosphorylation protection activities of AMP.

METHODS SUMMARY

Truncated AMPK (rat α and γ , and human β) was cloned into a tricistronic vector and expressed in *Escherichia coli*. Proteins were initially purified by nickel affinity chromatography. Crystals were grown by vapour diffusion in hanging drops using isopropanol as precipitant and were cryo-protected with 30% ethylene glycol. Crystals were soaked in ATP or Mg·ATP made up in reservoir solution. Diffraction data were collected at 100 K and processed using Denzo and Scalepack. The structures were solved by molecular replacement and refined using Refmac5 with manual model building using O. Bound nucleotide analysis was carried out by denaturation of samples, removal of protein and analysis by HPLC. *N*-methylanthraniloyl (mant) nucleotides were prepared according to published methods. Uncorrected fluorescence emission spectra of the mant nucleotides and their complexes with AMPK were recorded at 20 °C using a SPEX FluoroMax fluorimeter. Binding of mant-AXPs was monitored by titrating 10 μ M of the nucleotides with AMPK. Dissociation constants for AMP and ATP were determined using competition assays.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 April; accepted 9 August 2007.

Published online 12 September 2007.

1. Kahn, B. B., Alquier, T., Carling, D. & Hardie, D. G. AMP-activated protein kinase: Ancient energy gauge provides clues to modern understanding of metabolism. *Cell Metab.* 1, 15–25 (2005).
2. Hardie, D. G., Carling, D. & Carlson, M. The AMP-activated/SNF1 protein kinase subfamily: metabolic sensors of the eukaryotic cell? *Annu. Rev. Biochem.* 67, 821–855 (1998).

3. Hardie, D. G. & Carling, D. The AMP-activated protein kinase: Fuel gauge of the mammalian cell. *Eur. J. Biochem.* **246**, 259–273 (1997).
4. Yeh, L., Lee, K. & Kim, K. Regulation of rat liver acetyl-CoA carboxylase. Regulation of phosphorylation and inactivation of acetyl-CoA carboxylase by adenylate energy charge. *J. Biol. Chem.* **255**, 2308–2314 (1980).
5. Hardie, D. G., Carling, D. & Sim, A. T. R. The AMP-activated protein kinase—a multisubstrate regulator of lipid metabolism. *Trends Biochem. Sci.* **14**, 20–23 (1989).
6. Minokoshi, Y. *et al.* AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus. *Nature* **428**, 569–574 (2004).
7. Viollet, B. *et al.* The AMP-activated protein kinase $\alpha 2$ catalytic subunit controls whole body insulin sensitivity. *J. Clin. Invest.* **111**, 91–98 (2003).
8. Zong, H. *et al.* AMP kinase is required for mitochondrial biogenesis in skeletal muscle in response to chronic energy deprivation. *Proc. Natl Acad. Sci. USA* **99**, 15983–15987 (2002).
9. Minokoshi, Y. *et al.* Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *Nature* **415**, 339–343 (2002).
10. Yamauchi, T. *et al.* Adiponectin stimulates glucose utilization and fatty-acid oxidation by activating AMP-activated protein kinase. *Nature Med.* **8**, 1288–1295 (2002).
11. Watt, M. J. *et al.* CNTF reverses obesity-induced insulin resistance by activating skeletal muscle AMPK. *Nature Med.* **12**, 541–548 (2006).
12. Zhou, G. *et al.* Role of AMP-activated protein kinase in mechanism of metformin action. *J. Clin. Invest.* **108**, 1167–1174 (2001).
13. Shaw, R. J. *et al.* The kinase LKB1 mediates glucose homeostasis in liver and therapeutic effects of metformin. *Science* **310**, 1642–1646 (2005).
14. Sanders, M. J., Grondin, P. O., Hegarty, B. D., Snowden, M. A. & Carling, D. Investigating the mechanism for AMP activation of the AMP-activated protein kinase cascade. *Biochem. J.* **403**, 139–148 (2007).
15. Carling, D. The AMP-activated protein kinase cascade—a unifying system for energy control. *Trends Biochem. Sci.* **29**, 18–24 (2004).
16. Woods, A. *et al.* Characterization of AMP-activated protein kinase β subunit and γ subunit—assembly of the heterotrimeric complex *in vitro*. *J. Biol. Chem.* **271**, 10282–10290 (1996).
17. Hudson, E. R. *et al.* A novel domain in AMP-activated protein kinase causes glycogen storage bodies similar to those seen in hereditary cardiac arrhythmias. *Curr. Biol.* **13**, 861–866 (2003).
18. Iseli, T. J. *et al.* AMP-activated protein kinase β subunit tethers α and γ subunits by its C-terminal sequence (186–270). *J. Biol. Chem.* **280**, 13395–13400 (2005).
19. Townley, R. & Shapiro, L. Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. *Science* **315**, 1726–1729 (2007).
20. Rudolph, M. J., Amodeo, G. A., Bai, Y. & Tong, L. Crystal structure of the protein kinase domain of yeast AMP-activated protein kinase Snf1. *Biochem. Biophys. Res. Commun.* **337**, 1224–1228 (2005).
21. Corkey, B. E., Duszynski, J., Rich, T. L., Matschinsky, B. & Williamson, J. R. Regulation of free and bound magnesium in rat hepatocytes and isolated mitochondria. *J. Biol. Chem.* **261**, 2567–2574 (1986).
22. Scott, J. W. *et al.* CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. *J. Clin. Invest.* **113**, 274–284 (2004).
23. Chen, Z. *et al.* AMPK signaling in contracting human skeletal muscle: acetyl-CoA carboxylase and NO synthase phosphorylation. *Am. J. Physiol. Endocrinol. Metab.* **279**, 1202–1206 (2000).
24. Fujii, N. *et al.* Exercise induces isoform-specific increase in 5' AMP-activated protein kinase activity in human skeletal muscle. *Biochem. Biophys. Res. Commun.* **273**, 1150–1155 (2000).
25. McConell, G. K. *et al.* Short-term exercise training in humans reduces AMPK signalling during prolonged exercise independent of muscle glycogen. *J. Physiol.* **568**, 665–676 (2005).
26. Fryer, L. G., Parbu-Patel, A. & Carling, D. The anti-diabetic drugs rosiglitazone and metformin stimulate AMP-activated protein kinase through distinct signaling pathways. *J. Biol. Chem.* **277**, 25226–25232 (2002).
27. Gollob, M. H. *et al.* Identification of a gene responsible for familial Wolff–Parkinson–White syndrome. *N. Engl. J. Med.* **344**, 1823–1831 (2001).
28. Arad, M. *et al.* Constitutively active AMP kinase mutations cause glycogen storage disease mimicking hypertrophic cardiomyopathy. *J. Clin. Invest.* **109**, 357–362 (2002).
29. Blair, E., Redwood, C., Ashrafian, H., Ostman-Smith, I. & Watkins, H. Mutations in the $\gamma 2$ subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis. *Hum. Mol. Genet.* **10**, 1215–1220 (2001).
30. Adams, J. *et al.* Intrasteric control of AMPK via the $\gamma 1$ subunit AMP allosteric regulatory site. *Protein Sci.* **13**, 155–165 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Work in both laboratories was supported by the Medical Research Council (UK). D.C. acknowledges support from the EC. We are grateful to N. Justin and I. Taylor for MALS analysis, and S. Smerdon for discussions.

Author Information Coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 2V8Q, 2V92 and 2V9J. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.R.M. (smartin@nimr.mrc.ac.uk), D.C. (david.carling@csc.mrc.ac.uk) or S.J.G. (sgambli@nimr.mrc.ac.uk).

METHODS

Crystallography. Truncated AMPK, $\gamma 1(\text{rat})\beta 2(\text{human}; 187\text{--}272)\text{His-}\alpha 1(\text{rat}; 396\text{--}548)$, was cloned into a tricistronic vector and subsequently expressed in *E. coli* BL21 (BL21-CodonPlus-RIL, Stratagene)³¹. Proteins were purified using a nickel affinity chromatography (His-Trap, GE Healthcare), anion exchange (Mono Q, GE Healthcare) and gel filtration (Superdex 200, GE Healthcare). Protein complex stock solution was prepared at 15 mg ml^{-1} in 50 mM Tris, pH 7.0, 100 mM NaCl and 1 mM TCEP, and then mixed with a threefold molar excess of AMP. Crystals were grown by vapour diffusion technique at 18 °C in hanging drops. Drops were prepared by mixing equal volumes of protein complex with reservoir solution containing 15% isopropanol. Crystals were first transferred into mother liquor with an additional 30% ethylene glycol, before plunging into liquid nitrogen. For crystal soaking experiments, 1 mM ATP/0.1 mM EDTA and 0.5 mM ATP/0.6 mM MgCl_2 were made up in reservoir buffer, crystals were soaked for 1–5 h before flash cooling. Diffraction data were collected on an in-house MicroMax 007HF rotating anode coupled to a RaxisIV⁺⁺ detector. Data were integrated using Denzo and scaled with Scalepack³². The structure was solved by molecular replacement using Amore³³ using 2OOX.pdb as the search model, because of the differences in domain orientation between the yeast and mammalian enzyme the calculations were first performed with the γ domain, which was then fixed, and the search repeated with the combined α and β domains. Standard refinement was carried out with Refmac5 (ref. 34) and CNS³⁵ together with manual model building with O³⁶. Figures were created with Pymol (<http://pymol.sourceforge.net/>) and Grasp.

Bound nucleotide analysis. Solutions of known concentration were set to pH 2.0 to denature the protein, adjusted to pH 4.0 and then centrifuged to remove denatured protein. The supernatant was applied to a Partisil 10 SAX column (Whatman) and eluted with 0.6 M $\text{NH}_4\text{H}_2\text{PO}_4$ at 1.5 ml min^{-1} . The eluent was monitored for absorbance at 280 nm and the signal monitored by an integrator (see Supplementary Information).

Binding studies. *N*-methylanthraniloyl (mant)-nucleotides were synthesized, purified and characterized as described elsewhere³⁷. Uncorrected fluorescence emission spectra of the mant-nucleotides and their complexes with AMPK were recorded at 20 °C in 25 mM Tris, 1 mM TCEP (pH 8.0) using a SPEX FluoroMax fluorimeter with $\lambda_{\text{ex}} = 366$ or 380 nm (bandwidth 1.7 nm) and emission scanned from 400 to 550 nm (bandwidth 5 nm). Binding of mant-AXPs was monitored by titrating 10 μM of the nucleotides with AMPK. Dissociation constants for AMP and ATP were determined using two different types of competition assay (see Supplementary Information for further details).

31. Neumann, D., Woods, A., Carling, D., Wallimann, T. & Schlattner, U. Mammalian AMP-activated protein kinase: functional, heterotrimeric complexes by co-expression of subunits in *Escherichia coli*. *Protein Expr. Purif.* **30**, 230–237 (2003).
32. Otwinowski, Z. & Minor, W. in *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. & Bailey, S.) 556–562 (SERC Daresbury Laboratory, Warrington, 1993).
33. Navaza, J. AMoRe: an Automated Package for Molecular Replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
34. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
35. Brunger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
36. Jones, T. A., Zhou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
37. Jameson, D. M. & Eccleston, J. F. Fluorescent nucleotide analogs: synthesis and applications. *Methods Enzymol.* **278**, 363–390 (1997).

naturejobs

**JOBS OF
THE WEEK**

The demographics of academia in the United States are due for an overturn in the next decade. With the 'greying' of the universities — a third of US university faculty are aged 55 or older, and 14% are between 60 and 69 — institutions are preparing for a crew change, according to an article in the London *Times Higher Education Supplement* (14 September 2007).

The types of job — and the pay packets — the new crew are likely to get may, however, be very different from those commanded by the old guard. Many universities are anticipating replacing high-paying, full-time, tenure-track positions with lower-paying, part-time, adjunct ones. And the jobs will probably come as a trickle over the next decade, rather than a deluge, as the United States does not have a mandatory retirement age.

Nevertheless, it is good news for young scientists looking to enter US academia in the next few years, and not only for Americans. The United States has produced disproportionately fewer home-grown PhDs in the sciences — an area with one of the higher potential retirement rates — compared with other disciplines, and there is likely to be a demand for overseas scientists as well to fill the empty teaching positions.

The most coveted jobs released by retiring academics will undoubtedly be those at the nation's most prestigious institutions, such as the Massachusetts Institute of Technology, Harvard University and Columbia University, where a high proportion of faculty is older than 70. These universities are less likely than others to convert tenured posts into adjunct positions, as they have the endowments to fund professorships and reputations that would be eroded by replacing stars with part-timers.

Some universities are concerned that with the influx of young faculty, they won't have enough experienced people to manage departments and organize programmes. But young scientists who have done multiple postdocs and had the bar for advancement repeatedly raised are unlikely to be sympathetic. It's time to give them their chance.

Paul Smaglik was editor of Naturejobs from 2001 to 2007.

CONTACTS

Acting Editor: Gene Russo

European Head Office, London

The Macmillan Building,
4 Crinan Street,
London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:

Andy Douglas (4975)
e-mail: a.douglas@nature.com
**Business Development
Manager:**
Amelie Pequignot (4974)
e-mail: a.pequignot@nature.com

Natureevents:

Claudia Paulsen Young
(+44 (0) 20 7014 4015)
e-mail: c.paulsenyoung@nature.com

France/Switzerland/Belgium:

Muriel Lestringuez (4994)

Southwest UK/RoW:

Nils Moeller (4953)

Scandinavia/Spain/Portugal/Italy:

Evelina Rubio-Hakansson (4973)

Northeast UK/Ireland:

Matthew Ward (+44 (0) 20 7014 4059)

North Germany/The Netherlands:

Reya Silao (4970)

South Germany/Austria:

Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:

Stephen Russell
To send materials use London
address above.

Tel: +44 (0) 20 7843 4816

Fax: +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Jasmine Myer

US Head Office, New York

75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager:

Peter Bless

Japan Head Office, Tokyo

Chiyoda Building,
2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:

Ayako Watanabe
Tel: +81-3-3267-8765
e-mail: a.watanabe@natureasia.com

ENTER THE DRAGON

Once a poor village, Shenzhen is now one of the wealthiest cities in China. **David Cyranoski** learns its plans for the future.

Shenzhen is a city of transformations. Until 1979, it was a small fishing town in southern China. Then in 1980, Communist Party chairman Deng Xiaoping designated Shenzhen, located in Guangdong province just a few miles from Hong Kong, as a 'special economic zone', where tax policies encouraged foreign investment and trade. Since then the city has boomed, with its population growing to 12 million and fishing shacks being razed in favour of towering skyscrapers.

In what came to be known as 'the factory of the world', the city's economy soared to its current gross domestic product (GDP) of 581 billion yuan (US\$76 billion), the fourth largest in China. It produced clothing, shoes and toys quickly and cheaply, and assembled simple electronic goods.

A second transformation is now under way. With active support from local and central government, the city is metamorphosing from contract manufacturer to knowledge producer. The turning point came with the formation of a university town in 2003, comprising graduate schools from three of China's leading academic institutions far to the north: Peking University and Tsinghua University, both in Beijing, and Harbin Institute of Technology in Harbin.

Many academics refer to Shenzhen as 'special' or an 'experiment', Ma Hui, associate dean of the Graduate School at Shenzhen, Tsinghua University, is one. He calls the city both "a window and an open door": its proximity to Hong Kong (see map opposite) can provide a flow of free-market funding and its high profile can show the West how much and how fast China is changing. Just how good the picture from the window will be, and just how wide the door will open, could well depend on how the city manages a balancing act between free-market economy and top-down planning and between basic and applied research.

To create the human resources and technology



"This city has the entrepreneurial spirit to make things happen."

— David Ho

needed by a high-tech city, Shenzhen's policy-makers set about creating a combined science park and university town. Most scientific hubs are based on existing strong research universities, but Shenzhen had no such academic core. So the high-tech industrial park came first, in 1996, led in its early years by Liu Yingli, who subsequently became executive vice-mayor and a key force behind Shenzhen's economic development. Then, in 2001, the city persuaded the graduate schools to set up shop as University Town. Before this, there was just one university, Shenzhen University, which had around 25 graduate students. "That wasn't enough for a modern city," says Zhang Rulin, deputy director of University Town. Of the 4,500 students now studying in University Town, 800 are pursuing PhDs.

L. HUANG/Z. LIU

Coming home

Although Shenzhen's declared aim of becoming an international research centre is still a work in progress, increasing numbers of Chinese scientists are coming to Shenzhen from overseas. One is Zhang Yaou, a professor of cellular and molecular biology in the life-sciences division of Tsinghua University's Shenzhen campus. After earning a medical degree in China, she gained research expertise in oncology and haematology in Canada and the United States before moving to Hong Kong, while keeping an eye on China's political and intellectual environment. She's glad she made the move, noting that the Shenzhen students are of the highest quality and as bright as any she's worked with elsewhere.

To ensure these high standards, two-thirds of the teaching staff at the Tsinghua graduate school come from the home institution in Beijing, either residing in Shenzhen or travelling down periodically to teach courses or lead research. The other third mainly comprises returnees like Zhang Yaou. The researchers



The dragon-shaped library serves as a central point of the University Town campus.



come because they get access to new facilities, generous funding and a freer environment, in which they can adopt new interdisciplinary approaches to research. The graduate schools of Tsinghua University, Peking University and Harbin Institute of Technology were each awarded three 'key projects', each worth 30 million yuan, last year by the local Shenzhen government to establish nine multidisciplinary research clusters.

Shenzhen's development hasn't escaped high-profile attention in the West. David Ho, director of the Aaron Diamond AIDS Research Center in New York, is backing the chemical genomics laboratory at Peking University's Shenzhen graduate school by serving as chair of its advisory committee. The lab comprises biologists and chemists working together to identify biological problems and resolve them using a combination of biological and chemical techniques. The lab may help pave the way for biotech in Shenzhen, says Ho. The city is committing resources to facilities and equipment, but the biggest challenge will be attracting people with experience in the industry. Ho is optimistic. "This city has the entrepreneurial spirit to make things happen," he says.

All the Shenzhen institutes acknowledge the flexibility to restructure research projects, and even the buildings themselves, to allow interdisciplinary research. "This would not have been possible in Beijing," says Shi Shouxu, assistant president of Peking University and dean of the Shenzhen graduate school. "It's like a blank piece of paper. You can draw whatever you want," says Jin Guangjun, dean of the Harbin Institute of Technology Shenzhen graduate school.

And there are other advantages to Shenzhen. Being part of a special economic zone means that researchers have relatively easy access to isotopes, cells, proteins and other equipment that can get hung up in customs' red tape and demands for special permission elsewhere



Executive vice-mayor Liu Yingli (top) has helped make Shenzhen the place where Feng Guanping (bottom), director of the Research Institute of Tsinghua University in Shenzhen, has spent "the most exciting 10 years of my life".

in China. "It makes a world of difference when setting up large-scale screening," says Zhang Hui, a molecular biologist from Yale University who has taken up a post in the chemical genomics lab.

Ho's involvement in the Shenzhen project emphasizes what Zhang Baoquan, director of the Shenzhen Municipal Education Bureau, calls a focus on the elite. Bringing in top faculty from University Town's parent institutions ensures quality students. "We've had three classes of graduates, and the companies that hire them say they are of the same or better quality as those from the Beijing campus," he says.

The green space of the University Town campus contrasts with the skyscrapers and neon lights downtown. Symbolic of the campus's ambitious goals is the library in the shape of a dragon that runs through the centre of the site. Costing 220 million yuan to build, the library gets 20 million yuan per year for books, databases and periodicals.

Shenzhen speed

New research institutes have also been established in Shenzhen. The Research Institute of Tsinghua University Shenzhen, which opened in 1999, made news in the spring of 2003. The virus responsible for sudden acute respiratory syndrome (SARS) was creating a panic in Beijing, and Chinese president Hu Jintao asked Feng Guanping, the president of the institute, for ideas. Ten days later, Feng says, they had made an infrared sensor that when pointed at a person's head could detect the low-grade fever characteristic of SARS. The device soon became a fixture at airports. Feng describes his decade in Shenzhen as "the most exciting 10 years of my life." He adds "I wish I had come here earlier."

That episode became a symbol for what Deng early on had labelled "Shenzhen speed" — very fast development to meet social need. The establishment in 2006 of the Shenzhen Institute of Advanced Technology, affiliated with the prestigious Chinese Academy of Sciences, was another coup for the city. Focusing on biomedical and health engineering, robot services, high-performance



J. KANG

computing, and other interdisciplinary fields, the institute has 600 staff, with plans to grow to 3,000. The director, Fan Jianping, says his three-year term should have the institute stable. "In Shenzhen, if you can build a totally new city in ten years, and a university in six years, you can build an institute in three," he says.

Fan points out of a window in his 200-million-yuan building to a gray five-story building with grime-coated windows only a block away. "Last year this building was just like that," he says. "We are trying to make the blue-collar workers into white-collar workers."

Shenzhen is aiming its sights at converting applied research into saleable products. One of Fan's goals for his institute is patent generation, but the patents must be good. For every three patents applied for, a researcher will be expected to get at least one invention licensed.

Tsinghua's Feng comments that, "the big difference with the Beijing main campus is that they have many research achievements, but they don't go to the market. Here they will." Indeed, almost everything in Shenzhen is market-driven, both as an end — commercial products — and in organization and hiring policies.

The Shenzhen Virtual University Park provides continuing education as well as helping start-up businesses. Their curriculum is determined in "market oriented" fashion, says director Yi Yinfa. If the classes are not popular they are dropped. "In Shenzhen there is no 'iron rice bowl' that guarantees permanent employment," says Jin. "Everything is a two- or three-year contract. We want competition. We want people that are driven for success," he says.

Striking a balance

Basic science is not neglected, however. In collaboration with the National Orchid Conservation Center in Shenzhen, researchers at the Tsinghua University Shenzhen campus recently discovered and characterized an unknown type of self-pollination used by tree-living orchids in times of drought and lack of wind (Liu, K. W. *et al. Nature* **441**, 945; 2006; a video of the orchid can be seen at natureasia.com/Shenzhen).

Ma thinks that the balance between basic and applied science will change with time. "As we establish ourselves in Shenzhen, we will move more and more upstream towards basic research, always aiming to stay a step ahead of Shenzhen's rapidly evolving high-tech industry," he says. Research is already addressing real-world problems as well. For example, the Cooperative Research and Education Center for Environmental Technology at Shenzhen, funded and staffed jointly by Tsinghua University and Kyoto University in Japan, is looking at ways to provide cleaner water in urban areas.

And Shenzhen already has some role models for applied technology and industry. Telecoms companies Huawei Technologies and Zhongxing Telecom (ZTE) were founded in Shenzhen in the 1980s, and in 2006 brought in 66 billion yuan and 24 billion yuan in



Before and after: Shenzhen's landscape has changed dramatically since 1982 (inset).

"We are trying to make the blue-collar workers into white-collar workers."

— Fan Jianping

revenue, respectively. ZTE pours 11.6% and Huawei 8.2% of that back into research.

Such companies are helping Shenzhen do something rare for China's booming cities — they are ploughing huge amounts of money back into the local economy. In 2004, Shenzhen-based companies produced more than 50% of the city's GDP for the first time (by 2006 it was 58%). In places like Shanghai, 70% of GDP is produced by foreign investment companies.

"The high-tech industries of Shenzhen's regional innovation system are maturing, and this is setting a strong example for the whole country," says Zhang Jing'an, a member of the Party Leadership Group of the central government's Ministry of Science and Technology and president of the newspaper *Science and Technology Daily*

(www.stdaily.com). Indeed, governments in developing countries around the world that are trying to build their own special economic zones are looking to Shenzhen as a model.

The city has reinvested its wealth with some headline-catching results. The gene-therapy company SiBiono was set up in Shenzhen in 1998 and had its treatment approved by the Chinese regulatory authorities for head-and-neck squamous cell carcinoma in 2003 (*Nature Biotechnol.* **22**, 3–4; 2004). Some 6,000 people have now been treated. Peng Zhaohui, the company's founder and chairman, says he chose Shenzhen over Shanghai and Beijing because of the financial support offered by the local government.

There are plenty of people to fill spots at these companies — for every successful applicant at Huawei, 99 don't get in. But not enough of them have the right talents. Peng says, "It's hard to get high-level staff. There were no famous universities." Peng ends up recruiting 5% of his staff from overseas and most of the rest from other parts of China.

The growth of the three universities in University Town will help provide more local talent. The Hong Kong and Shenzhen local governments also agreed this May on greater collaboration in R&D in various areas, including health and environmental protection. "Hong Kong has advantages in attracting international talent and connecting with the international markets," says Zhang Jing'an. "Shenzhen has advantages in developing innovative industries and gathering mainland talent."

But to attract and keep its creative talent, Shenzhen needs to allow room for basic research and intellectual pursuits that don't have immediate economic returns. Given that governments everywhere are increasingly demanding a return on their research investment, Shenzhen's strategy is not so extreme. Still, it remains to be seen whether it can spur the type of Silicon Valley innovation that Shenzhen seeks. But if it can be done, Shenzhen will do it — and in a hurry.

David Cyranoski is Nature's Asia-Pacific correspondent.

naturejobs

Correction

In the Regions feature 'Enter the dragon' (*Nature* **449**, 502–504; 2007), we quoted Yi Yinfa as saying that the curriculum at the Shenzhen Virtual University Park is determined in a "market oriented" fashion. This quote should in fact have been attributed to the university park's director, Qiu Xuan.

NATUREJOBS PRESENTS THE HIGHLIGHT ON IRELAND

DATE: 25 OCTOBER 2007

DEADLINE FOR ADVERTISING: 19 OCTOBER 2007

Naturejobs is pleased to present the Highlight on Ireland, appearing in the 25 October issue of *Nature*. If you're looking to recruit top-tier scientific talent to your organisation or raise your company profile amongst the European scientific community, you won't want to miss this issue.

This Highlight offers a unique opportunity for companies and organisations in Ireland to profile themselves and publicise their collaborations and activities to a high-quality audience.

If you'd like to learn more about advertising in the Highlight on Ireland, please contact one of our representatives.

Don't delay - the deadline for advertising is this issue is 19 October 2007.

T: +44 (0)20 7014 4059

E: m.ward@nature.com

If you are interested in advertising events in Ireland please call +44 (0)20 7014 4015 or e-mail c.paulsenyoung@nature.com

CALL TODAY TO ADVERTISE WITH NATUREJOBS!

WWW.NATUREJOBS.COM/CONTACTUS

naturejobs

nature publishing group **npg**

IN112580R



SPOTLIGHT ON NORTHERN VIRGINIA AND MARYLAND

DATE:

OCTOBER 18, 2007

ADVERTISING DEADLINE:

OCTOBER 11, 2007

Northern Virginia and Maryland are both home to numerous top-tier universities and major research institutions. Recent government allocations have boosted funding and furthered development of these areas as scientific hubs. *Naturejobs*, the career and recruitment magazine in *Nature*, will present an advertising platform for organizations to give job seekers a closer look into these important states.

A dedicated section of *Naturejobs* will include recruitment, event, and scientific announcements advertisements highlighting the region. The Spotlight precedes the regular recruitment section and be marked to draw readers' attention.

Key benefits:

- Attract top scientists from industry and academia to your company/organization
- A strong association is made between your organization and the states of Virginia and Maryland
- Your job or event ad will be posted to our online database free for 8 weeks to www.naturejobs.com
- This Spotlight will have prominent positioning on our *Naturejobs* home page, permanently archiving your advertisement, and exposing the region even further

Nature is read by over 616,000 readers each week, and online *Naturejobs* receives over 2.1 million page views each month.

FOR MORE INFORMATION, PLEASE CONTACT YOUR NATUREJOBS REPRESENTATIVE:

Virginia:

Leslie Pineda

T: 212-726-9395

E: l.pineda@natureny.com

Maryland:

Shelley Cohen

T: 212-726-9229

E: s.cohen@natureny.com

www.naturejobs.com

naturejobs

nature publishing group **npg**



Shenzhen is indeed a pot of gold at the end of the rainbow.

Few, if any, cities in the world can boast the kind of dramatic change that Shenzhen has experienced. A fishing village in 1980, it is now a roaring metropolis. In a city of immigrants from all over China who are brimming with confidence and ambition, municipal organizers are converting its economic base from low-end manufacturing to high-tech industry. Expanding ties with neighbouring Hong Kong symbolize the city's prodigious emergence onto the world stage and promise a new chapter in this extraordinary story of growth and development. Shenzhen is a model for governments in developing countries around the world that wish to spur regional economic growth.

Until the late 1970s, Shenzhen was, like the rest of China, part of a strictly planned economy. Then, Deng Xiaoping, reinstated around that time as the head of the Communist Party, launched a series of reforms aimed at opening up the country to more economic growth, especially through ties and trade with foreign countries.

As part of these reforms, he opened up a handful of 'special economic zones', which were able to give tax incentives for foreign-owned companies. They could also open up certain aspects of the market to international trade. Results came immediately and foreign investment poured in.

Shenzhen, in the southern part of the southern province of Guangdong, was the first city granted these privileges. Zhuhai and Shantou, also in Guangdong, Xiamen, in the neighbouring Fujian province, and the island of Hainan also became special economic zones. But none of these took off like Shenzhen. In 1992, Shenzhen's economy was on the same scale as the other four special economic zones, but now its total economy exceeds all of them added together. Now it is six times greater. At RMB 581.36 billion (USD 77.25 billion), its gross domestic product is the fourth highest in China.

Immigrants, ideas and innovation

Zhou Luming, the deputy director general of the city's Bureau of Science, Technology and Information, says one of the advantages

was simply the proximity to Hong Kong, a well-developed centre of international finance. But the main reason was the flood of people who came from all over China, bringing their hopes and dreams with them. Among the city's 12 million inhabitants, only 400,000 are native.

"The passion of people who had been restrained for many years under the planned economy system exploded in this city," says Zhou.

"The immigrants brought new ideas and innovation," adds Lin Xiong, deputy director of the bureau's Planning and Development Department.

During the 1980s, the city made its fortune on low-end manufacturing, as foreign companies were looking for cheap places to manufacture their goods. Shenzhen's economy was growing by 50% a year.

In the 1990s, city officials decided to upgrade their economy by establishing a high-technology infrastructure. They had been spurred by competition from the other special economic zones and a natural desire to produce more profitable, higher-end products. They also now had the economic wherewithal to invest in such technology.

And they had models near at hand. Two telecommunications companies, Huawei Technologies Co., Ltd. and Zhongxing Telecommunication Equipment Co., Ltd. (ZTE), founded in 1988 and 1985, respectively, took off in the early 1990s to become two of the world's largest companies in the field today. The twins of Shenzhen, as Huawei and ZTE have come to be known, together account for some RMB 90 billion in industrial productivity, over 8% of the city's total. ZTE is now rushing to catch up with Huawei. International sales increased nearly 100% over the first half of 2007. The two companies are now creating R&D centres throughout China, as well as overseas in places like India. "When Shenzhen was looking for innovation, these two companies grabbed the chance," says Lin.

From 1992 to 2006, the average annual growth of Shenzhen's high-tech industries increased by a staggering 45%. In 2006, Shenzhen's high-tech production reached RMB 630 billion,

Index

- 1-2 From village to high-tech powerhouse
- 3 A man on a mission
- 4 A platform for growth: the China Hi-Tech Fair
- 5 Shenzhen Virtual University Park
- 6 Graduate School at Shenzhen,
Tsinghua University
- 7 Peking University Shenzhen Graduate School
- 8 Harbin Institute of Technology Shenzhen Graduate School

- 9 Shenzhen University
- 10 Huawei Technologies
- 11 Zhongxing Telecommunication
Equipment (ZTE) Corporation
- 12 BYD Company Ltd.
- 13 Tencent Holdings Ltd.
- 14 SiBiono GeneTech Co., Ltd.
- 15 Netac Technology Co., Ltd.
- 16 Shenzhen Chipscreen Biosciences Ltd.
- 17 China International Marine Container (Group) Ltd.

Publisher information

Publisher: David Swinbanks
Sales director: Kate Yoneyama
Project manager: Ayako Watanabe
Project supervisor: Jeff Kang
Local coordinator: Felix Cheung
Correspondents: Adarsh Sandhu, John Fox
Editors: Robin Taylor, Linda Worland, Meg Fitzpatrick
Production: Takesh Murakami, Seiko Makita

topping China's other major cities. The city's investment in research and development is some 3.4% of its GDP — far higher than the national percentage of 1.4%.

Unlike most technological innovation that remained in the laboratory in other cities, the market economy system, established by the local government in Shenzhen, saw enterprise there take a dominant role in technological innovation and achieve great commercial success. Despite the blueprint for Shenzhen's high-tech industry being considered a joke 10 years ago, state leaders now laud Shenzhen as a model for other cities across China to follow in developing their high-tech industries.

In contrast to other cities, Shenzhen has adopted a 'reverse innovation model', in which technological innovation is driven by market demand. Enterprises orientate the direction of innovation according to market demand, and then develop their innovations through research collaborations with universities. Using this model, Shenzhen's enterprises have progressed from low-end to high-end innovation.

As an administrator of the city, Shenzhen government plays its role well to meet the demand of the market economy, and creates a relaxed environment to develop enterprises. And as a city that has legislative power, Shenzhen was the first in China to introduce legislation in 1995 that protects intellectual property rights — 'Regulations of Shenzhen Special Economic Zone on the Protection of Technical Secrets of Enterprises'. This made Shenzhen the city that best protects intellectual property rights in China.

But Shenzhen still had a challenging task — to train and establish the required personnel to complete the transformation to a high-tech economy. At that point, the local government believed it was a suitable time to establish its own universities and research institutes.

Fast and efficient solutions

"In order to meet the demands of a city, in terms of culture and education, we had built Shenzhen University and two vocational schools. They had succeeded, but they could not meet the demands of the rapidly growing city," says Zhang Baoquan, director of the Shenzhen Municipal Education Bureau.

Shenzhen's solution to this problem was characteristic of what came to be known as 'Shenzhen speed' and 'Shenzhen efficiency'. Faced with a shortage of high-tech graduates, Shenzhen's leaders decided to invite some of the country's premier research institutions, Peking University, Tsinghua University, Harbin Institute of Technology, and Nankai University, to establish branches. Instead of aiming for new schools that could feed education to the masses like other university towns, they gambled on elite education. Zhou says that Shenzhen's decision to attract universities in the late 1990s was well timed with the decision by those institutions to become more market-orientated. Shenzhen's proven ability to bring products to market made it an attractive partner.

Shenzhen also established many research institutes, such as the Research Institute of Tsinghua University in Shenzhen in 1999, and the Shenzhen Institute of Advanced Science and Technology of the Chinese Academy of Sciences in 2006. It also opened up a 'Virtual University Park' in 1999 to help the universities and other research institutions develop and capitalize on their technological prowess.

This top-down approach, importing high-technology research and education, was unlike anything seen in China, where the emphasis is usually on slow 'normal' development. "This might seem to be against Chinese philosophy, but we dare to do things first," says Zhang.

Zhang says there have been sceptics. "Many people think we could not keep up the standards of these prestigious universities, but history shows we can," he says, referring to the formation of the University of Cambridge by scholars leaving the University of Oxford.

The city has continued to invest in these universities, and this year they have each received three key laboratories, worth RMB 30 million each. In 2006, Shenzhen had 29,000 patents, the second highest number of any city in China. Now Shenzhen has the confidence to build its own comprehensive science and technology university, and construction will soon begin on an RMB 4 billion campus whose doors are set to open in 2010.

Integrating and enhancing assets

The research and human resources in the making, as well as the active investment environment, are likely to attract more companies. The Shenzhen Stock Exchange, the first and still one of only two in China, will be reborn by the end of this year to become a Nasdaq-style exchange that encourages more high-technology start-ups, allowing companies to post shares after just two years of profitability. The move is expected to encourage hundreds of small companies to list.

Shenzhen continues to impress. On a visit in 2005, Premier Wen Jiabao stated that Shenzhen should be a role model for other cities in China.

Policymakers in Hong Kong have also recognized the city's achievements, and this will lead to an entirely new phase in Shenzhen's remarkable growth. In the past, Hong Kong has resisted efforts to forge tighter links with its mainland neighbour. But on May 21, Joseph Wong, Hong Kong's secretary for commerce, industry and technology, and Liu Yingli, Shenzhen's executive vice mayor signed an agreement on technological collaboration called the 'Shenzhen-Hong Kong Innovation Circle'.

The idea behind the collaboration is to integrate the complementary assets of each: Hong Kong's well-established universities and researchers, experience with intellectual property, and international finance know-how with Shenzhen's wealth of entrepreneurs, manufacturing ability, innovation and sheer determination.

Executive Vice Mayor Liu says that the Hong Kong collaboration is the key to the city's future growth. Even greater ties might be in the offing. The idea of a Hong Kong-Shenzhen metropolis was mapped out in a recent report by Bauhinia Foundation Research Centre, a Hong Kong-based think tank closely tied to the government. The report found that facilitating transport and travel between the two cities and increasing financial co-operation would be mutually beneficial and would create a Hong Kong-Shenzhen metropolis, with an area greater than London, to become the third largest in the world by 2020.



Shenzhen's vibrant research environment and the availability of human resources, as well as the active investment environment, is attracting many companies to the city.

A man on a mission



Executive Vice Mayor Liu Yingli of Shenzhen Municipal Government.

Liu Yingli, executive vice mayor of Shenzhen and one of the key innovators behind the development of the city, is a man on a mission.

“Shenzhen is often misunderstood,” he says, while explaining the city’s rapid development.

We are more than a commercial city providing a factory at the back of the Hong Kong shop-front, he explains. From the beginning, Shenzhen planned a high-tech industry and a university.

Liu cites four major factors in Shenzhen’s success. First is the opening up of Shenzhen as a special economic zone by the central government. “That allowed us to take the lead in developing a market economy, which is a fundamental reason for our success”, he says.

The second factor, he believes, was timing. At that time, in the early 1980s, there was an industrial shift in which multinationals were beginning to outsource to centres around the world and Shenzhen offered a good low-cost option.

“Third,” says Liu, “citizens, government officials and academics came from all corners of China to Shenzhen. These people were by nature enterprising, innovative and ambitious. Innovation is in their blood,” he enthuses.

The fourth factor Liu points to is proximity to Hong Kong. “Hong Kong has a well-developed market economy. It’s a world-class financial, trading and logistics centre and we benefit a lot from those attributes.”

Liu sees Hong Kong as key to the future development of

Shenzhen. He is the prime instigator on Shenzhen’s side of the ‘Shenzhen-Hong Kong Innovation Circle’, a joint initiative of the two local governments to create a ‘united metropolis’.

Four universities in Hong Kong — Hong Kong University of Science and Technology, the University of Hong Kong, City University of Hong Kong, and the Chinese University of Hong Kong — are each putting HK\$50–100 million (US\$6.5–13 million) per year into research centres in Shenzhen.

The relationship between Hong Kong and Shenzhen is very different compared with 20 years ago and the presidents of Hong Kong universities know mainland China well, according to Liu.

Twenty-eight academics of the Chinese Academy of Sciences are in Hong Kong, and Hong Kong universities now have many students from the mainland. These universities will play a “prime role” in the Innovation Circle, explains Liu.

Equally important are the academic institutions in Shenzhen, where Tsinghua University has had a presence for more than 10 years, first as a research institute and now as a graduate school. One hundred and thirty universities have carried out projects in Shenzhen according to Liu.

So, what are the key challenges that lie ahead for Shenzhen? There are many, according to the vice mayor. Foremost is the recruitment of talented people, who, he says, are “worth a fortune” for science and technology development.

“The flow of talented people is without boundaries in this global world,” notes Liu. “One of our biggest challenges will be to attract and keep good people.”

Liu Yingli — Biography

From farm worker to executive vice mayor of Shenzhen, Liu Yingli’s career had an unusual start.

Born in 1948, Liu was sent to a farm in Jilin Province as an “educated youth” from 1968 to 1970, during the Cultural Revolution. Then, until 1978, he was group leader of the Planning and Management Group of the First Automobile Works in Changchun, Jilin Province. Liu was one of the “lucky class of ’77” allowed to re-enter university when China re-started higher education. He studied applied mechanics at Jilin Industry University and stayed to complete a masters degree in computational mechanics of engineering technology. Then, Liu took the unusual step of accepting a job as an engineer in the Electronics Re-

search Institute of Shenzhen, when the city was little more than a small fishing village. During this time he spent a year in Canada working for a software company, ACDS.

In 1992, Liu joined the local Shenzhen government and rapidly rose up the ranks of the Financial Planning Department of the Bureau of Science and Technology to eventually become head of the department. He was then appointed assistant to the general director of the Bureau of Science and Technology of Shenzhen before becoming vice director of the bureau. Since then he has held many appointments, including vice secretary general of the Shenzhen Municipal Office from 1996 to 2002, and director of the Lead Group of the High-Tech Industrial Park of Shenzhen city since August 1997.

From September 2002 to May 2005, Liu was vice mayor of the Shenzhen Municipal Government and a member of the municipal Communist Party committee. From January 2003, he was also the director and secretary of the Communist Party committee of the Lead Group of the High-Tech Industrial Park, chairman of the Research Institute of Tsinghua University in Shenzhen, chairman of the National Engineering Centre for Biotechnology in Shenzhen, and dean of the Chinese Academy of Science and Technology Development. From May 2005, he has been a standing member of the municipal Communist Party committee and executive vice mayor of Shenzhen Municipal Government, while continuing in his many other roles.

A platform for growth: the China Hi-Tech Fair



www.chtf.com/english

The China Hi-Tech Fair was established in 1999 as a platform for international companies, investors, inventors, entrepreneurs and relevant government bodies to meet, exchange ideas, make business deals and establish collaborations.

Organized by the Shenzhen Municipal Government, approved by China's State Council and co-hosted by eight government ministries and agencies, the Chinese Academy of Sciences, Chinese Academy of Engineering and the Shenzhen Government, the fair is held on October 12-17 every year.

Its huge success in just 9 years has attracted the world's attention, putting Shenzhen on the high-technology map.

"The city needed to create its own 'name card' to create a brand name," says Wang Xuewei, director general of the Shenzhen Bureau of Trade and Industry who supervises the fair.

In this city of immigrants, Wang's 20 years in Shenzhen make him one of the most experienced policymakers. Formerly an engineer at the Institute of Optics, Fine Mechanics and Physics at the Chinese Academy of Sciences, Wang watched during the late 1980s as companies from the US and Europe moved in to manufacture clothing, shoes and toys. He watched during the 1990s as companies from the same developed countries began to produce and assemble basic electronic goods such as televisions and radios. Shenzhen companies were usually the original equipment manufacturer of products that were repackaged and sold by the foreign companies.

But he also watched as companies in Shenzhen and throughout China gained competency in these areas of basic electronics, setting the stage for China's emergence in the high-tech industry. "Now we need a platform to develop these strengths," he says.

The China Hi-Tech Fair is doing the job. "We want to use this event to enhance our production base in these areas," says Wang. "It is a bridge between technology resources and enterprises, and its purpose is to show the world what Shenzhen wants to be."

Companies and regional governments from all over China attend the fair, but Shenzhen has a special place because of its proven manufacturing ability, which promises to give life to bold ideas. It is a natural complement for many of the companies or entrepreneurs at the fair. "People come here with ideas and designs, but no ability to produce," Wang says. "Here we may not have all the ideas, but we know we have the ability to produce."

The fair has grown rapidly as word has spread about the opportunities it presents. Between 1999 and 2006, the number of countries and regions represented grew from 26 to 42. The number of programs more than doubled, from 4,150 to 9,765, as the number of exhibitors increased from 2,856 to 3,278. The number of investors jumped from 955 to 2,690. Twenty-five foreign governments send delegations, compared with five in



Boundless opportunities for anyone with a good idea.



Opening a new chapter in Shenzhen's history.

1999. And the number of visitors has doubled, from 300,000 in 1999 to 612,000 in 2006.

The fair leads China in focusing on topics of vital concern to the country. In 2006, for example, the fair hosted a 'Recycling Economy Exhibition'. The exhibition's 7,500 square metres showcased leading technologies in energy conservation, new energy, ecological and environmental protection, population and health, and green agriculture. This tribute to sustainable development, a topic of crucial importance to China, was capped with a special section devoted to the Bluesky Award for Renewable Energy Technology and Investment. The special event was hosted by the United Nations Industrial Development Organization.

Last year's fair also featured another topic that has been in the headlines as China emerges on the world's high-technology stage — intellectual property rights. The event featured a forum with senior intellectual property experts from companies and law firms, from China and abroad, to discuss the special challenges faced in China.

Wang relates a story about someone who came to the first fair with patents but little else and, during the course of successive fairs, made connections that positioned him as the chief executive officer of a RMB 100 million company. Wang says there have been many such success stories, and the people of Shenzhen have so much energy that he expects to see many more. "People who think they have something great see an opportunity here and so they come," he says.

Filling a gap in the innovation system



Shenzhen Virtual University Park

The Shenzhen Virtual University Park was established in September 1999 at a crucial time in Shenzhen's development.

In 1980, Shenzhen became the first of Deng Xiaoping's 'special economic zones'. The following 10 years marked Shenzhen's emergence as a cheap manufacturer of products outsourced from Western countries. During the 1990s, Shenzhen embarked on a plan to develop its own core technology, to train and attract more technologically advanced human resources, and to generate intellectual property.

Qiu Xuan, the director of the Shenzhen Virtual University Park says, "As the competition between Shenzhen's processing enterprises intensifies, their individual expectation to take the lead in technological innovation and to own core technology creates market demand for high-end technologies. Usually you would look to the universities to take the lead in innovation, but there was not enough university-based research and training here. Because it was such a young city, we had to look outside."

The Virtual University Park is devoted to bringing in high-calibre researchers, teachers, and intellectual property to help the city's development. Forty-eight research institutions have joined the park, including the University of Hong Kong, the Chinese University of Hong Kong, Hong Kong University of Science and Technology, and three other top Hong Kong institutions. Several institutes, among them the prestigious Chinese Academy of Sciences, the Chinese Academy of Engineering, and the Chinese Academy of Social Sciences, also joined. Top Chinese institutions such as Peking University and Tsinghua University signed up, as did four overseas universities, including the Lyon Central Polytechnic University in France, the University of Alberta in Canada, Budapest University of Technology and Economics, and the University of Miskolc in Hungary.

The high-tech enterprises in Shenzhen have now expanded and urgently need to export their products and trade technologies. This forced them to cooperate with foreign universities. Qiu says that the Virtual University Park is looking to recruit high-end talent, industrialize achievements, and develop international exchanges. At the moment, foreign universities are actively being introduced.

The park serves two main functions. The first is to provide higher education, especially targeted at workers in Shenzhen who want to improve their skills or attain a higher degree, and join the city's high-technology boom. Many of these students are sent by their employers for additional training.

Despite its name, the university provides face-to-face courses. Professors from 35 of the member universities travel to the park to give courses, often on weekends or evenings. Administrators carry-out market research with the universities to determine which courses will be needed. If they are well attended, they survive, says Qiu. A logistics course offered by the Shanghai University of Communications is so popular that only one-third of the students can get in each term, says Qiu. The park has trained 70,000 students, and bestowed nearly



Shenzhen Virtual University Park provides higher education to Shenzhen's workers who want to join the city's high-technology boom.

1,000 doctorates and 17,000 masters degrees.

The park's main function, however, is technology transfer and the licensing of technology from its member universities and other research organizations. An area of 160,000 square metres is devoted to incubating enterprises. Another 180,000 square metres of office space for incubators is under construction.

"More than 200 companies have been born from our incubators using technology from member universities," says Qiu. "Of the incubated projects, we have a 90% success rate."

To help businesses grow, the park organizes meetings among local investment companies, government agencies involved in the promotion of industry, and professional associations. The Virtual University Park is also building its own university science park, a project that has been granted National University Park for Science and Technology status by the Ministry of Science and Technology and the Ministry of Education.

Qiu's next goal is to bring some State Key Laboratories to the park. China has about 180 of these prestigious institutions, mainly within the universities and national laboratories. They each support a few hundred graduate students. Qiu says this would help the park carry out its role of filling the gaps in Shenzhen's emergence as a high-technology power. "Shenzhen is short on research, just like it was short on higher level university training," says Qiu. "Our most important work is to build the laboratories that can give local industry support."

Shenzhen Virtual University Park
www.szvup.com

Bridging academic excellence and innovation with entrepreneurship



清华大学 深圳研究生院
Graduate School at Shenzhen, Tsinghua University

Graduate School at Shenzhen, Tsinghua University

'One university, one brand' is a guiding principle of Tsinghua University's new graduate school in Shenzhen. It reflects the school's intention to maintain the high standards of academic excellence of its home institution Tsinghua University, far to the north in Beijing. At the same time, the Shenzhen graduate school seeks to create a new culture through interaction with Shenzhen's vibrant economy and industry.

Tsinghua University ranks as one of China's top universities, alongside Peking University, its neighbour in Beijing. It is quite a coup for the Shenzhen local government to have persuaded both of these prestigious institutions to set up graduate schools in Shenzhen's new university town.

"Our graduate school is focused on developing professionals with leadership skills, international vision, and an enterprising and innovative spirit in a multidisciplinary research environment," explains Guan Zhicheng, the school's dean.

Established on a new campus on the outskirts of Shenzhen only 4 years ago, the school has about 2,000 graduate students, with nearly 1,500 of them based full-time in Shenzhen. All students do some course work at the Beijing campus. About 80 full-time faculty and 200 part-time faculty from Beijing are spread across five divisions: information science and technology, engineering, life science, management, and science and liberal arts.

Two-thirds of the full-time faculty are either from Tsinghua University or have connections to Tsinghua.

"We apply exactly the same procedures for hiring faculty for Shenzhen as the main campus and therefore standards are just as high," explains Guan. Similarly, students face the same rigorous standards for entrance to the graduate school and for completion of degrees as at the main campus.

"I was not sure if I wanted to come to Shenzhen but once I got here I really liked it," says Guan. The campus, next to Lake Xili and the mountains, about 30 minutes drive from the centre of Shenzhen, provides a quiet environment for study. The rapidly expanding high-tech industry of Shenzhen is nearby and interactions with industry are actively encouraged.

Tsinghua University, which is particularly strong in engineering



The Tsinghua University Graduate School campus, next to Lake Xili and the mountains, provides a quiet environment for study.

but is also branching out into life science and medicine, generates about 800 patent applications a year. The Shenzhen graduate school has already produced 82 in its short life.

The school aims to be a conduit for the transfer of technology from Tsinghua University to industry. As an example, Guan points to technology for a production line for compact disks that was initially developed at the main campus but has been taken forward at Shenzhen with 14 additional patents. This production line has now been lent to a local company for further development. Until now, such production lines have tended to be imported from overseas.

Professors in the Management Division have also taken part in developing an important new road map for an IP strategy for both the local and central government. Others are advising on transport infrastructure and logistics.

The first batches of graduate students have completed their degrees. While the majority have gone elsewhere for employment, a significant proportion of them — particularly law students — have found jobs in Shenzhen.

The Shenzhen local government is clearly pleased with the graduate school and is giving it RMB 90 million in one-off funds to set up new cluster laboratories. Tsinghua is establishing laboratories in three areas: 'clean production' to help local industry reduce pollution and increase efficiency; 'new media technology', including 3D video systems; and 'new medicines' where the focus will be on basic research. With funding from Guangdong Province the school has also set up a key laboratory for research in chemical biology.



A production line for compact disks, developed initially at the Beijing campus, has been taken forward for further development at the Shenzhen campus.



Guan Zhicheng,
Dean of the Graduate School at Shenzhen,
Tsinghua University.

Tsinghua University Graduate School
Tsinghua Campus, University Town, Shenzhen 518055, P.R. China
www.sz.tsinghua.edu.cn/en

An interdisciplinary new generation



北京大学深圳研究生院
Shenzhen Graduate School
PEKING UNIVERSITY

Peking University Shenzhen Graduate School

The challenge for the Shenzhen Graduate School of Peking University was living up to the reputation of its prestigious main campus in Beijing while carving out its own identity as a dynamic research institution that will help the development of one of the fastest growing cities in the world. After 5 years of innovative policies and generous support, the school is succeeding on both fronts.

The graduate school keeps close links to its home institution in Beijing. Students must complete the same entrance procedures, and faculty must meet the same rigorous hiring process. Faculty are largely recruited from Peking University or overseas.

The new graduate school has one definite advantage over Beijing. Old academic barriers do not prevent interdisciplinary studies. Environmental scientists and urban development specialists work together easily, while information technology is taking off on the basis of common ground shared by scientists and engineers working on software, systems, and semiconductor chips.

One of the best examples of interdisciplinary studies is the Laboratory of Chemical Genomics. With more than 1,800 lines of transgenic zebrafish and two common-use nuclear magnetic resonance machines, chemists and life scientists work together in the same building. Biologists identify and characterize important molecular interactions, and chemists design molecules needed to study them. Completing the loop, biologists feed the molecules they have created back into cellular systems for testing. The laboratory, with renowned HIV researcher David Ho as the head of its scientific committee, aims to provide targets for treating HIV, HCV and cancer.

Ye Tao, a member of the chemical genomics laboratory who received his PhD at Queen's University in Belfast, says the laboratory not only exemplifies the type of interdisciplinary research possible in Shenzhen, but is also unique in being run by a group of scientists who all trained in the West and work in a manner similar to Western universities.

The involvement of Ye, a professor at Hong Kong Polytechnic University, also symbolizes the growing network of Hong Kong scientists who are bringing their expertise and entrepreneurial spirit to take advantage of the opportunities in Shenzhen. More of this interaction will occur since the two cities are in the process of finalizing a united 'Hong Kong-Shenzhen metropolis' plan.

Yale University's Zhang Hui, who also leads a research group in the chemical genomics laboratory, says that multinational companies are rushing in. "They want to come to China, and they look around and find our lab."

Wu Yundong, a member of the Chinese Academy of Science and also the vice chairman of the scientific committee of the chemical genomics laboratory, emphasized the laboratory's unique operating system. The scientific committee works closely with the laboratory's director, Yang Zhen, to ensure that interdisciplinary research work meets milestone-based targets and the goals for strategic research and development of the laboratory. The director and the principal investigator's



Research at the Peking University Graduate School complements the strengths of Shenzhen's industry (left).

The chemical genomics laboratory, where much interdisciplinary research takes place (right).

committee are responsible for all the hiring, resource allocation and appraisal of laboratory members.

The graduate school will also devote itself to supporting the municipality's development. Law, business and social sciences curricula produce new talent for various sectors of government and private industry. In natural science and engineering, the graduate school has chosen to focus on fields such as information technology and life sciences that "match the strengths of Shenzhen industry", says Executive Dean of Peking University Shenzhen Graduate School, Shi Shouxu.

Such research will help the city reach its dream of becoming a high-tech hub, while at the same time ensuring that fundamental science underlying the necessary innovation is considered. "We won't give up basic research but we will focus on research that might produce benefit for Shenzhen," Shi says.

After 5 years, the school now has 1,811 students. Thirty per cent of graduates stay in Shenzhen taking up positions in law or at one of the booming electronics companies like Huawei or ZTE.

With generous funding provided equally by the local government, the central government, and industry, further growth is certain.



Executive Dean of
Peking University Shenzhen Graduate School,
Shi Shouxu.

Peking University Shenzhen Graduate School
Lishui Road, NanShan District, Shenzhen 518055, P.R. China
www.szpku.edu.cn/ips

A free environment and a window of opportunity for science and technology

The Harbin Institute of Technology (HIT) is 3,034 kilometres from its new graduate education and research arm, the HIT Shenzhen Graduate School. But the distance between the famous mother institution in China's northeast and Shenzhen on the southeast coast means little, says Jin Guangjun, the school's executive dean. "HIT is not geographically limited. It is a brand name," he says.

And it is a powerful brand name. Founded in 1920, HIT has developed into one of China's strongest institutions in computer science, semiconductor and laser engineering, materials science, and mechanical engineering including robotics. In 1999 it was selected as one of nine Chinese scientific institutions to receive additional funding from the government so that it could develop its potential to world-class levels.

The new graduate school will benefit from HIT's expertise, experience, and history, but HIT also stands to gain from the fresh ideas that are starting to come from the south, ideas that provide a window to the future. With some 2,000 universities in China all fighting to show their value, HIT has to stay young, fresh, and well connected to industry needs. The Shenzhen campus, in the bustling and expanding economic hub of China's Pearl River Delta, is helping it meet that goal. "We want to keep our top-nine status," says Jin.

Jin says he was attracted to Shenzhen because of the city's spirit of independence: "We have freedom to do what we want, which is something different from the traditional Chinese university." In Jin's own field of urban design, the expanding city allows him to let his imagination unfold and take advantage of multidisciplinary expertise at the school. Over the past three decades Shenzhen has been transformed from a fishing village to a massively crowded and complex city. "The basic architecture is done. Now we have to make it a more comfortable place," says Jin, who has been working on the 'New City Center' project in the Bao'an district.

While maintaining strong ties to its home institution, the Shenzhen Graduate School will have a great degree of independence. The school can open up new interdisciplinary fields of endeavour, which will in turn give it the freedom to work flexibly with international collaborators and with industry.

Indeed, promoting industrial collaboration is a driving force behind the graduate school. Already there are five laboratories from which HIT staff collaborate with local Shenzhen companies. Local industry funds some 65 other projects with RMB 16.4 million.

It is rapidly becoming one of China's most international schools, with 50% of courses taught in English. To achieve this, the school has taken advantage of increasingly strong ties with neighbouring Hong Kong. More than 30 professors from Hong Kong universities have adjunct positions at the graduate school. Some are paid, while others come to take advantage of the research funding available in Shenzhen. "Many more Hong Kong-based professors are interested in coming," says Jin. The



Harbin Institute of Technology Shenzhen Graduate School



HIT staff from the graduate school's five laboratories collaborate with local Shenzhen companies.

broad range of English curricula will help the school appeal to prospective foreign staff and students. "We want to broaden the international horizon. We would love to see even more foreigners here," he says.

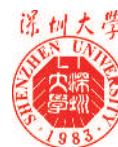
The number of articles published by graduate school staff in key scientific and engineering journals, as indexed by the Institute of Scientific Information, jumped about tenfold between 2002 and 2006. Since its first group of graduate students signed up in June 2002, the school now has some 1,500 masters students and more than 300 doctoral students from all over China. So far, graduates have had a 100% employment rate, with 70% being hired by Shenzhen's booming companies. The very practical training given at the school has been key to this success, which benefits HIT, the graduate school, the students, and Shenzhen.



Executive Dean of
HIT Shenzhen Graduate School,
Jin Guangjun.

Harbin Institute of Technology Shenzhen Graduate School
HIT Campus Shenzhen University Town Xili,
Shenzhen 518055, P.R. China
www.hitsz.edu.cn

Human resources fuel economic transformation



Shenzhen University

In early 1983, Deng Xiaoping, the reforming head of the Chinese Communist Party, was driving down the road near Shenzhen's Back Bay when he passed a huge construction site. He was told it was the site for Shenzhen University and although construction had only started in spring, students would start enrolling the following fall. Deng nodded with a smile and said, "Shenzhen speed."

The university has become an engine for the spectacular growth of Shenzhen where, on Deng's suggestion, a special economic zone had been established a few years earlier.

"The history of the university is the history of the city," says Zhang Bigong, the university's president. Indeed, many of the buildings that have come to symbolize Shenzhen's soaring economic growth were designed in the university's school of architecture.

Unlike most universities in China that are supported by the central government, Shenzhen University is funded by the city government. This gives it a greater degree of flexibility, which its leaders have used to establish a pioneering institution.

Shenzhen University's first two goals were to provide human resources and technology for the special economic zone; and to contribute to internationalization of the city by providing a forum where Chinese and foreign researchers could exchange ideas.

Finally, the university itself would be an experiment in higher education. It has fulfilled this goal with policies such as putting its faculty on short-term contracts rather than hiring them as permanent staff, charging students a fee to cover part of their tuition costs, and giving a committee of professors the final say in important matters instead of the university president.

Such pioneering efforts have made Shenzhen University a "hot" institution in just over 20 years, says Zhang.

Although it was a new university, Deng's support enabled Shenzhen to recruit high-quality faculty, mostly from Peking or Tsinghua universities. The 1,800 faculty include many notable researchers, such as Li Jingzhen, who has pushed the frontiers of extremely high-speed photography; and Ruan Shuangchen who has made lasers with practical applications in spectroscopy, fiber optics, medicine, and display. Niu Zenqiang of the university's Optoelectronics Institute created a YAG laser that has been commercialized by Shenzhen United Winners



Shenzhen University campus.

Laser System. Sales of the lasers, in Japan and elsewhere, have already topped RMB 9 million. Liu Zhigang has capitalized on research into common allergens to produce an allergen diagnosis protein chip, which has been commercialized by Shenzhen Synogene Digital Co., Ltd.

Many of the university's 27,000 graduates have also become key players in Shenzhen's economic development. Li Yi, a graduate in electronic precision machinery, has become the chair of the board of Shenzhen TRONY Science and Development Co., Ltd. Shi Yuzhu, who completed a masters degree in soft science management at Shenzhen University in 1989, went on to start a software company (Giant Corporation) and was, for a time, China's eighth richest man.

The university's most famous success story is Ma Huateng. After graduating with a bachelors degree in computer science in 1993, Ma went on to co-found Shenzhen Tencent Computer System in 1998.

The university's contribution to Shenzhen has earned it ongoing support, including RMB 90 million for scientific research in 2006 from government and local industry, nearly a 50% jump over the previous year. The university is now finalizing plans for a medical school, funded with RMB 3 billion.

Internationally, the university's presence is being felt, with 600 international students from 40 countries and 30 of its faculty from overseas. Zhang has a vision of making Shenzhen University one of China's top 20 universities over the next decade. The success achieved already and continuing support makes him confident it can be done.



University President,
Zhang Bigong.

Shenzhen University
Nanhai Avenue 3688, Shenzhen 518060, P.R. China
www.szu.edu.cn/en



Shenzhen University has high hopes that its work in photonics will open up new industry in Shenzhen.

Chinese culture and Western expertise prove a successful mix



Huawei Technologies

Manufacture of telecoms equipment and provision of next-generation solutions for fixed, mobile and data communications are core business for Huawei Technologies.

More than one billion users and 31 of the world's top 50 telecommunications operators — including British Telecom, Vodafone and Telefonía — use Huawei's products. From a small trading agent to a global telecommunications giant, the story of Huawei, the most-respected enterprise of Shenzhen, reveals how the market-orientated high-tech companies in Shenzhen were driven by customer demand, and how they developed and grew through innovation in management and technology.

"Understanding and communicating with our customers is an important factor for our success," says Xu Zhijun, Huawei's chief marketing officer.

"Since 1996 we have consulted with companies, including IBM and Hay Group, to develop a management style based on Western expertise which includes Chinese culture. Our management process enables us to speak the same business language as our customers."

Huawei also continues to innovate based on customer requirements and, unlike many other vendors, continued to invest in R&D during the 'IT bubble'. In fact, 48% of staff are involved in R&D, which gives the company a leading edge in terms of both technology and solutions. The 3G base station is one such solution.

"Conventional 3G base stations are 'cabinets' which require a lot of space which the operator must rent," says Xu. "The initial cost is only about US\$10,000–20,000 but the operating costs are high because the operator has to rent the space and pay for the large power supply.

"We provide distributed NodeBs, which use less space and power with the result that running costs are cut by 30%. European companies, including Vodafone and Orange, are big users of this distributed NodeB solution," he says.



Staff dormitory and swimming pool; Huawei offers competitive packages to attract talented staff.

Huawei offers very competitive packages to attract talented staff. "It was difficult in the early days but our brand is very well known now and it is easier to find good people. Now for every one hundred CVs we accept only one," says Xu. Huawei senior managers are aged in their 30s and typically manage businesses of US\$1–2 billion.

Huawei Technologies was founded in 1988 and is wholly owned by its 70,000 employees. The company started in Shenzhen from a small base and has grown with the support of Shenzhen's pro-business policies.

Challenges for the company have included the transition from domestic to international markets. In 1998, when Xu managed the company's international business, the perception of Chinese products was shoes, toys and low-quality products.

"When I talked with customers they did not believe that a Chinese company could produce high-tech equipment."

Huawei made a resolute decision to go global.

"We were determined to expand to overseas markets, which is how we are here today," explains Xu. Now, over 65% of sales are overseas and in 2006 the company generated net revenues of US\$8.5 billion and US\$11 billion in contracts.

But, going global required an internationally recognized management process, which the company found through consulting with many Western companies. Future expansion is planned to take advantage of China's low R&D costs.

"Our business model of high-quality, excellent service, low operating costs and giving priority to customer requests cannot be copied by Western companies," says Xu.

Intellectual property is high on the list of priorities for Huawei, which has submitted more than 19,000 patent applications to date. "Our company is only 19 years old and we are willing to discuss cross-licensing. We are prepared to buy patents. That is our strategy," he says.

How far can Huawei go in the future? Xu is confident, "We want to become the Toyota of telecoms! We are now aiming for this target."



Huawei Technologies data centre.

Huawei Technologies

Bantian, Longgang District, Shenzhen 518129, P.R. China

www.huawei.com

Global success from local wisdom



Zhongxing Telecommunication Equipment (ZTE) Corporation

ZTE Corporation is the world's fastest growing manufacturer of telecommunications equipment and wireless solutions. The company, which was founded 22 years ago by Hou Weigui, is China's only telecoms company with shares traded publicly on both the Hong Kong and Shenzhen stock exchanges.

An ongoing investment in R&D has enabled ZTE to provide customized technologies and services to companies in more than 120 countries. It is the only Chinese telecoms manufacturer to be listed in *BusinessWeek's* 2005 Top 100 Information Technology Companies, as well as *BusinessWeek's* 2006 Top 20 Brands.

But how did a small start-up evolve into a US\$4 billion company providing 3G technology after only 21 years?

"Our success is firstly tied to China's open-door policy launched in the early '80s, which stimulated a huge demand for telecommunications-related equipment," says senior vice president, Xie Daxiong.

"Shenzhen's designation as a special economic zone enabled us to grow initially with production of digital switches, then fixed transmission and mobile communications."

ZTE now has more than 100 products, embracing wireless, wireline, service and terminal technologies.

"Shenzhen has always been very helpful to us," says Xie.

"Our initial market was rural China, then we concentrated on the cities and finally we went global. Now, 52% of our sales are from overseas business." In the first half of 2007, ZTE's total international business revenue was RMB 7.97 billion. This figure has increased by 99% in comparison with the same period in 2006.

Forty per cent of the company's 40,000 employees are involved in R&D, across 15 wholly owned R&D centres in North America, Europe and Asia. The R&D staff are hired locally and given well-defined objectives.

"We recruit the best people to achieve our goals," says Xie. "From the early days of launching the company, we have invested about 10–12% of our annual revenue on R&D."

ZTE also provides customers with training at nine dedicated facilities worldwide, including the ZTE University, which has trained more than 130,000 people in China and overseas.

At the end of July 2007, the company had submitted more than 9,300 patents. Attaining intellectual property rights has been a core element of its strategy since inception.

"We produce our own IPR portfolio which enables us to interact with much larger companies through cross-licensing agreements," explains Xie.

Examples of ZTE's patented innovations include the world's first CDMA mobile telephone with a detachable SIM card, which was launched in 2000 and the world's first global open trunking architecture system (GoTa) CDMA-based digital technology announced in 2004. These technologies were the first to be licensed to overseas vendors by a Chinese telecoms manufacturer.

"The future is in wireless communications; both systems and handsets," says Xie. "Chip design is one of our core technologies which we will use for development of new products and solutions."

Success is based on motivated and innovative staff, and ZTE recruits from China's top universities. Employees have an average age of 30, about 70% are university graduates and, of these, 25% have a masters degree and 1.4% have a PhD. ZTE has been voted 'Most Favorable Enterprise for Employment' for four consecutive years by 60 universities in China.

The company has a global presence, with collaborators that include Vodafone, Alcatel, Ericsson and France Telecom in Europe. It maintains three R&D centres in the US; has customers in South America including Brasil Telecom and VIVO; enabled the first 3G call made in Africa; and has customers in Saudi Arabia, Pakistan, Tajikistan, Mongolia and Vietnam. This global success epitomizes ZTE's motto of 'Global success from local wisdom'.



ZTE Senior Vice President,
Xie Daxiong.

ZTE Corporation

ZTE Corporation

ZTE Plaza, Keji Road South, High-Tech Industrial Park,
Shenzhen 518057, P.R. China
www.zte.com.cn

ZTE Corporation



ZTE handsets of the future.



ZTE headquarters and R&D building in Shenzhen.

Building dreams for the future



BYD Company Ltd.

The world's largest manufacturer of rechargeable batteries has successfully turned its hand to other IT components and automobiles. BYD Company Ltd., established by Wang Chuan-fu and 30 like-minded people in 1995, now employs more than 100,000 people and is listed on the Hong Kong stock exchange. This is a success story of Shenzhen's high-tech enterprises winning wealth through the integration of labour and technology, and ultimately walking the path of high-end research and development.

With the goal 'Build Your Dreams', president Wang says the founders took advantage of China's unique business factors — namely, the availability of low-cost, talented people and an enormous market.

"The huge demand for IT products and automobiles in China brought many overseas companies here," says Wang. "With them, they brought their supply chain and hence a route for our products. Our business model could not possibly be used in countries like the USA or Japan."

In the early years the company focused on the manufacture of rechargeable batteries. Intense investment in innovative R&D enabled BYD to produce high-quality, low-cost products that attracted customers including Motorola and Nokia.

The revenue generated from batteries and other IT components gives BYD the scope to enter other business sectors.

"We found that the battery and IT components market was difficult to expand," says Wang. "We decided to enter the automobile business because we thought it had a brighter future." In 2003 BYD bought the Tsinchuan Automobile Company, and created the BYD Auto Company Ltd.

"The battery and auto businesses are both labor intensive so we thought we could translate our experience [in] manufacturing high-quality batteries at competitive rates to the auto industry," explains Wang.

R&D is central for both businesses and BYD acquired non-patented technology to improve the quality of its products. The company has seven industrial bases across China, covering a network of R&D, production and logistics. The factory in Xi'an can produce 200,000 cars per year. The BYD F8 is the first car in China with a retractable glass hard-top roof.

Another important decision was locating the company's headquarters in Shenzhen. "In 1995 Shenzhen's special economic zone was very open to doing business. The municipal

government had many useful regulations for establishing businesses here. It was an age of the market economy and Shenzhen took the lead," says Wang.

The company places a high priority on recruitment, with production workers and R&D staff recruited from across China. R&D on materials, software and vehicle engineering is carried out at the BYD Central Research Institute and the Telecom and Electronics Research Institute, where activities include WiFi, acoustics and auto electronics.

Technicians and production staff are trained at BYD's technical school, which covers an area of 60,000 square metres within the BYD industrial park in Longgang. BYD provides free dormitories and sports facilities for its employees. They even have their own village with a kindergarten, club, restaurant and supermarket. The company fully sponsors its own BYD middle school, which is authorized by the Shenzhen Municipal Government.

Maintenance of intellectual property rights (IPR) is important for such a young company. "With 80–90% of our products for export, we must maintain a high-quality IPR portfolio. In 2006, the company was granted 1,166 patent applications." Altogether, BYD has 2,500 patent applications, including those granted and pending. More than 1,000 cases are filed in China and at least 50 cases in other countries each year.

BYD was chosen as one of the 'Stars of Asia' by *BusinessWeek* magazine in 2003 owing to the success of its Li-ion battery.

"The Li-ion business was dominated by Japanese industry but our products gave the market an alternative manufacturer," notes Wang.

The high quality and low cost of BYD's batteries have been key factors in their international acceptance.

Plans for the future are based on BYD's core technologies of IT components and automobile manufacture. The company is already using its batteries to produce electric cars.

"The F3e is our first battery-powered vehicle. BYD stands for 'Build Your Dreams' and we are confident that we can succeed with electric vehicles," says Wang.

BYD Company Ltd.

Yan'an Road, Kuichong, Longgang, Shenzhen 518119, P.R. China

www.byd.com

BYD Company Ltd.



The BYD F3e integrates BYD's IT and auto businesses.



BYD's journey began with its high-quality, low-cost rechargeable batteries. BYD's headquarters in Shenzhen.



BYD's president, Wang Chuan-fu, accepts *BusinessWeek* 'Star of Asia' award from former US president, Bill Clinton (right).

Multi-dimensional platform connects people across China

Tencent 腾讯

Tencent Holdings Ltd.

Tencent is China's leading internet company, connecting more than 600 million registered instant messaging (IM) user accounts and over 30 million peak simultaneous online user accounts. The company was launched in 1998 by Ma Huateng (CEO) and Zhang Zhidong (CTO), both computer engineering graduates from Shenzhen University. Tencent is a success story of Shenzhen local university students and how they combined IT and lifestyle into not only a new business model, but also the 'Shenzhen dream'.

"In 1998 the founders wanted to start a Chinese paging network, which we eventually called QQ, for connecting people in real time via a platform that used the internet," says the president, Martin Lau, who has a masters degree in electrical engineering from Stanford University and an MBA from the Kellogg Graduate School of Management, Northwestern University.

Tencent's QQ instant messaging platform was extremely successful, making a profit of US\$53.9 million in 2004. In June 2004, the company went public on the Hong Kong stock exchange where the initial offering raised HK\$1,787.8 million.

The QQ platform has grown tremendously because it provides a multi-dimensional way for people to stay connected. "QQ is not just text but a very rich means of staying connected either visually, using PC-to-PC instant messaging, as personalized animated characters called avatars, or by interactive games such as chess and cards," says Lau.

But how has Tencent succeeded in China where giants like Google and Yahoo struggled? The answer lies in China's unique internet landscape where more than half the internet users are under 30 and interested in using it for chatting and entertainment, rather than searching for information.

"QQ is successful because in China there is a huge demand for the internet as a means of socializing. Also, China is a huge country and people on the move want to stay connected so they use QQ," explains Lau.

A browse through Tencent's website also shows the variety of ways that people can connect. "We have content that enables a rich way of connecting. It's not just text messages as in e-mail. People can connect by playing games or sharing photo albums and music," says Lau.

Locating the company in Shenzhen was an important component of its success. "The availability of talented staff was extremely important," explains Lau. "Shenzhen attracts talented people from all over China because of its vibrant environment and entrepreneurial spirit.

"We could say that the city itself is a 'start-up' and its government entrepreneurial. The city government is proactive in a number of ways, such as giving tax advantages, the use of telecommunications infrastructure and organizing conventions to attract investors," he says.

Commenting on Tencent's recent award of 'Best Chinese Lifestyle Brand', Lau notes it shows that QQ has become a part



People can connect to Tencent's QQ by playing games or sharing photo albums and music.

of people's lives.

And for the future? "We are focusing on communications and community; entertainment, such as online games and videos; content in the form of music portals and video channels; and e-commerce — which is a new business area for us," says Lau.

"The internet has an enabling power. It is still in its infancy and many more innovations and business opportunities are yet to come."

Lau says that for Tencent to evolve, the key factors will be to attract motivated people, stay entrepreneurial and build around the company's core technologies and ideas.

It seems as though Tencent's mission to create an 'online society' in China has only just begun.



Tencent president, Martin Lau, says Shenzhen attracts talented people from all over China.

Tencent Holdings Ltd.

Fiyta building, Gaoxinnanyi Avenue, Southern District of Hi-Tech Park, Shenzhen 518057, P.R. China
www.tencent.com/index_e.shtml

Tencent Holdings Ltd.

Shenzhen base for world's first gene therapy



SiBiono GeneTech Co., Ltd.

Gendicine, a recombinant adenoviral-p53 agent primarily for head and neck cancer, was the world's first gene therapy product to be approved for commercial production when it was given the go ahead by China's State Food and Drug Administration (SFDA) in January 2004. That it was developed in Shenzhen reflects the Shenzhen economic zone's high-technology research and development strategy.

Gendicine was the brainchild of Peng Zhaohui, chairman of Shenzhen SiBiono GeneTech Co., Ltd. Gendicine received new drug certificate approval from the SFDA in October 2003, a manufacturing licence in January 2004 and has been on the market in China since April 2004.

A biochemist by training, Peng was previously a professor and director of the Institute of Molecular Biology at China's Nanfang Medical University. He has studied at Chiba University in Japan and at the University of California at Los Angeles and worked for a biotech company in the US, before returning to China in 1998.

After returning to China to establish SiBiono, Peng initially considered moving to Beijing or Shanghai, before finally opting to locate in Shenzhen because the Shenzhen government supported the project as part of its policy of attracting innovative high-tech companies.

"Another advantage of Shenzhen is the presence of a number of excellent universities," says Peng.

"Given the need for specialized staff in such a venture, as well as graduates from Shenzhen and elsewhere in China, SiBiono recruits specialists from overseas," he adds.

About 10% of SiBiono's staff are Chinese who have returned from overseas. The company provides competitive salaries and challenging career opportunities to help minimize staff turnover.

As of July 2007, Gendicine had been administered to more than 6,000 patients, about 400 of whom were Caucasian. Most treatments were for solid tumours such as head and neck squamous cell carcinoma (HNSCC), hepatocellular carcinoma (HCC), nasopharyngeal carcinoma (NPC), and other cancers

that are relatively common among Chinese populations.

"Gendicine has been used off-label for more than 50 different types of cancer, but the only approved indications are for NPC and HNSCC," explains Peng.

While Gendicine monotherapy is effective, it has significant synergistic effects in combination with chemotherapy, radiotherapy, surgery, or thermotherapy. For example, in a phase II/III clinical trial in HNSCC patients, 64% were in complete regression after 8 weekly injections of Gendicine combined with radiotherapy, which is 3 times greater than with radiotherapy alone. Five-year follow-up results demonstrated that Gendicine improved radiotherapeutic tumour control by 11 times versus radiotherapy alone in NPC patients, with the 5-year overall survival and tumour-free survival rates being 13.5% and 28% higher than the radiotherapy control group, respectively. In late-stage HCC, Gendicine with chemotherapy was associated with a short-term efficacy of 67%, which was 16% higher than chemotherapy alone, while the 1-year survival rate increased 19.4% with the combination.

Gendicine is safe, with no evidence of the extreme immune responses reported with adenoviral gene therapies elsewhere. "In about half of the patients there was a low-grade fever about 3 hours after the injection, which lasted 2 to 6 hours then spontaneously regressed," says Peng.

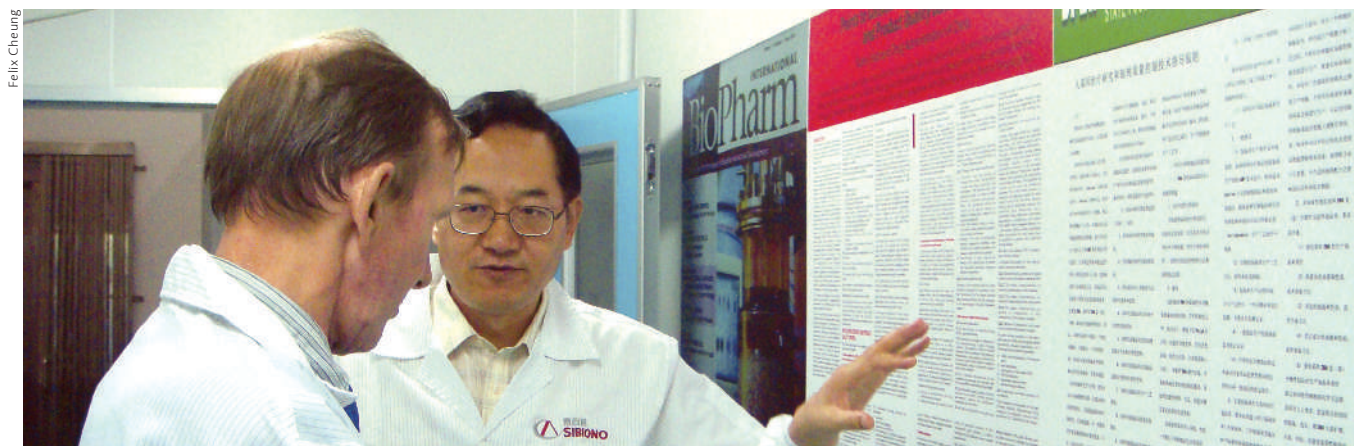
"At present, SiBiono's manufacturing capacity is about 150,000 Gendicine doses per year, which is likely to increase next year after the company moves to its new RMB 200-million production facility, still in Shenzhen. Gendicine is manufactured according to current Good Manufacturing Practice, a system for ensuring that medicines are consistently produced and controlled according to quality standards.

Shenzhen SiBiono GeneTech Co., Ltd.

Shenzhen Hi-Tech Industrial Park (North), Langshan Road,

Shenzhen 518057, P.R. China

www.sibiono.com



Chairman of Shenzhen SiBiono GeneTech Co., Ltd., Peng Zhaohui, was attracted to Shenzhen by government support for innovative high-tech industry.

Felix Cheung

Pioneer in mobile memory storage is Shenzhen success story

Netac

Netac Technology Co., Ltd.

Netac Technology Co., Ltd. illustrates the success of Chinese students who established their business in Shenzhen and ventured into the world via the China Hi-Tech Fair platform. Since its establishment in the Shenzhen economic zone in 1999 as a pioneer of USB flash drive memory storage systems, Netac has emerged as an R&D leader in mobile memory storage and digital entertainment products worldwide.

In 1999, while Frank Deng was still a student in Singapore, he heard that Shenzhen was organizing the first China High-Tech Fair. He immediately rushed to the city with his rough sample of the now ubiquitous 'flash drive' and an invention patent.

Deng attended the fair with almost no expectations. But, courtesy of the fair, he had access to a free exhibition booth, food and accommodation. And when an investor from Singapore heard of his invention, he contacted Deng immediately and decided to invest in developing the technology. Shenzhen government later provided RMB 12 million as venture capital funds so that he could file the patent for his invention.

"Shenzhen has excellent government services. Both the local and central governments encourage innovative technology companies, as China moves towards high-technology production and innovation," explains Netac's president and CEO Deng.

Netac's major activity in Shenzhen concerns R&D, which is carried out at the company's state-of-the-art facility in the Incubator Building of the Chinese Academy of Science and Technology Development. However, R&D can be prohibitively expensive, especially for a technology start-up company.

"The Shenzhen government has set up an IT design platform, which we use at a low cost," says Deng. "Such platforms are normally very expensive and the Shenzhen government allows all of the local IT design companies to rent the platform very cheaply.

"The science and technology department of the Shenzhen government has also been very helpful and supportive towards us and very quick in responding to any R&D or IT problems," he explains. Shenzhen seems to be better in this regard than other major cities of China, according to Deng.

Another advantage of Shenzhen is the abundance of qualified staff. "It is not difficult for us to recruit talented people," says Deng. While some of Netac's staff training is done in-house the company also sponsors some of their R&D and IT design engineers to do graduate studies at universities in Hong Kong.

"Of course, many of our engineers have also studied here in Shenzhen, where many of China's leading universities, including Tsinghua, Peking and Harbin, have branches. The Shenzhen government has attracted leading mainland universities to establish graduate schools in Shenzhen. This is a local government policy and a lot of our staff have studied there."

Netac offers expertise and services to develop products according to clients' needs. The company has a small manufacturing facility in Shenzhen and provides original design manufacture and original equipment manufacture to customers such



Netac is among the world's leading suppliers of mobile memory storage and digital entertainment products.

as Dell, IBM, Samsung and Toshiba. In fact, 80–90% of products are made on an 'original equipment manufacture' basis.

"We also sell our own brands to them," says Deng, adding that 70% of his company's output is for domestic consumption, with the rest being middle- to high-end items for export to markets worldwide.

Netac is among the world's leading suppliers of digital entertainment products, with a major focus on MP3 players for in-car audio systems.

"We are focusing particularly on USB hard drives. This is a very big market, not only for PCs, but also for use in TV," adds Deng.

Netac's only competition comes from manufacturers in the US and Taiwan. "We are absolutely number one in mainland China, where no-one can compete with us," says Deng, noting that Netac has filed more than 300 patents covering USB flash drive technology worldwide, including the US, Korea, Singapore, Hong Kong, China, Europe and Japan.



Netac President and CEO, Frank Deng.

Netac Technology Co., Ltd.
6F, Incubator Building,
Chinese Academy of Science and Technology Development
No.1 High-Tech South Street, Shenzhen 518057, P.R. China
www.netac.com

From biochips to drug discovery



Shenzhen Chipscreen Biosciences Ltd.

A biotech/pharmaceutical company specializing in drug discovery and development is unusual in China, where most pharmaceutical companies make generics. Shenzhen Chipscreen Biosciences Ltd., using its own integrated chemical genomics discovery platform, is one of the few companies operating in this area.

Chipscreen was founded in 2001 as a Sino-US joint venture by the company's director, president and chief scientific officer, Lu Xianping, and other returnees from the US.

"The first thing you notice about Shenzhen after living in the US is that it is very beautiful, is much cleaner and has better weather than other Chinese cities," says Lu.

"More importantly, several years ago Shenzhen began undergoing dynamic commercial development and we liked the strong market economy, which is really pushing us forward," he says.

In some ways Shenzhen reminds Lu of his former base in California. "I spent 10 years in San Diego, which, like Shenzhen, has an immigrant culture and is very vibrant and innovative."

The Shenzhen government was very supportive in terms of issuing permits and other paperwork needed to start the company and import the high-tech equipment they needed. Chipscreen also received about \$US2.5 million in government research grants.

"If we had tried to do this elsewhere in China, I'm not sure it would have worked," says Lu. Chipscreen is partnered with Tsinghua University, one of the company's largest shareholders and owners of the building housing its headquarters.

The company has more than 50 employees, 27% of them with PhDs. Investors come from Singapore, Hong Kong and from other Chinese cities.

"Most people working here have been here for several years, so we do not have problems retaining staff," says Lu, who plans to hire more people as the company continues to grow.

"Most of our PhDs are trained in the US, which is important," he says, noting that in the past, the Chinese generics-based industry had not been highly skilled in terms of innovation and discovery, although this is changing with the new generation of Chinese PhDs. At Chipscreen, "we are innovators, not imitators".

Chipscreen minimizes drug development time and costs by using its proprietary state-of-the-art chemical genomics technology to match potential small molecules to disease targets. Reference compounds are then used to screen for potential pharmacological activity and toxicities, which enables informed decisions to be made at an earlier stage in drug discovery development. In this way "we can decrease attrition rates and increase success rates", explains Lu.

The company has completed phase IIa clinical trials of its anti-diabetes drug chiglitazar on more than 240 diabetes patients and is about to enter Phase IIb. Regarding safety, "so far so good", he says, with no sign of the cardiovascular toxicity that has caused other agents in the same class to be withdrawn late in development. Upon successful completion of life-long toxicity studies in rodents, chiglitazar will be submitted for Phase III trials. If everything goes as planned, chiglitazar may



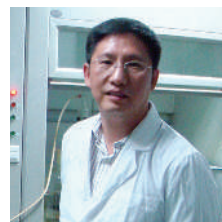
Chipscreen is partnered with Tsinghua University, one of the company's largest shareholders and owners of the building housing its headquarters.

be ready for product launch by 2010, according to Lu.

Chipscreen also has a new anticancer drug, chidamide, which is looking safe and effective in an ongoing Phase I trial.

"It is going extremely well, so last year we out-licensed the rights to chidamide, excluding use in China, to a US biotech company."

Developments in the pipeline include some exciting new molecules, including a series of enzyme inhibitors with potential in oncology and in areas other than cancer. After further development, the most promising will be contracted out to Chinese clinical research organizations for drug trials. Another product in development is an interesting compound derived from traditional Chinese medicine that appears to protect the liver and has received market authorization approval.



Chipscreen Director,
President and Chief Scientific Officer,
Lu Xianping.

Shenzhen Chipscreen Biosciences Ltd.

Suite C301, Research Institute of Tsinghua University,
High-Tech Park, Nanshan District, Shenzhen 518057, P.R. China
www.chipscreen.com

Global reach from Shenzhen



China International Marine Containers (Group) Ltd.

With more than 1,000 products in 10 categories, China International Marine Containers Ltd. (CIMC) is the world's largest manufacturer of shipping containers.

Most of the products are used in marine and railway transportation, including containers, vehicles and tank equipment. The rapid growth of CIMC is a stellar example of how high-tech enterprises in Shenzhen are relying on technological innovation to add value to their traditional products.

CIMC's growth mirrors that of its home city, Shenzhen, which over the past 25 years has grown from a cheap mass-production centre for Hong Kong factory owners to become the booming modern heart of China's high-technology sector.

When the company was first established in Shenzhen in 1980, the same year the city was designated a special economic zone, it was one of China's first Sino-foreign joint ventures. At that time, CIMC had an annual production of just 10,000 dry freight containers and was competing with about 20 similar companies in China. After listing on the Shenzhen Stock Exchange in 1994, CIMC expanded rapidly, building facilities at major ports along the China coast that gave it a competitive advantage over its rivals. CIMC has been the global leader in terms of volume since 1996.

CIMC chose Shenzhen as a base for commercial reasons. The city's status as China's first special economic zone was important for imports and exports.

"As a logistics company, we have to span the globe, and Shenzhen was the most convenient city in this regard," says Liu Chunfeng, the general manager of the R&D management division.

"When Shenzhen was first designated a special economic zone we had special tax breaks on exports, and exemptions for imports of raw material," Liu explains.

However, with China joining the World Trade Organization, he says that, in future, they will need to rely on the strength of their industry to compete.

"After more than 20 years of development, Shenzhen has



Zhang Liding, CIMC

CIMC manufactures more than 1,000 products used in global transportation.

come up to the standards of other international cities, so we are quite strong, especially compared with other cities in China."

In the area of staff recruitment, Liu says that, among talented people in China, the flow is now relatively free. "If the conditions you offer are very good, then you will attract the best talent in China, for example graduates from Peking, Shanghai Xiatong and Tsinghua universities." CIMC also hires internationally for specialized skills.

"For example, in the area of quality assurance, standards must be up to international levels and the talent in this field is from Europe or the USA," says Liu.

While most of CIMC's products, such as tanks and vehicles, are traditional, the company has been developing a Smart and Security Container (SSC), which uses a computer system and radio-frequency identification to track containers as they move around the world.

"A smart chip, including WiFi technology, can track the container throughout the whole transportation process," explains Liu.

"This will tell you the container's contents and whether it has been tampered with illegally. The initial aim of the SSC was for anti-terrorism; however, we realized that it also needed to be profitable, so we linked it to the commercial and logistics side."

As well as marine containers, CIMC has an airport equipment division that manufactures automatic air cargo and logistics handling systems, automatic parking systems, and passenger boarding bridges. These are used at more than 70 airports worldwide and include the boarding bridge for the new Airbus A380, the world's largest aircraft. "We are the first choice for most global transport businesses," concludes Liu.

China International Marine Containers Ltd.
No.2 Gangwan Avenue, Shekou Industrial Zone,
Shenzhen 518067, P.R. China
www.cimc.com



Zhang Liding, CIMC

CIMC has developed a system to track containers as they are transported around the world.

Alloy

An identity crisis.

Marissa Lingen

Humans have told my story for centuries: the artificial woman. Flowers, flesh or metal, my body fascinates. It is lush and unyielding, vibrant and sterile.

No one thinks of making an ugly woman from scratch. We have too many of the old-fashioned kind, I suppose — but the same is true of men, and Frankenstein's monster is no cover model. If I had been ugly, there would have been less trouble. But if I had been ugly, my maker would have taken out my bones and started again.

My maker was always practical.

In the stories, the creator made the artificial woman for himself. My maker dreamed bigger.

"You'll change the world," he told me, when I was too young to really converse. "You'll make everything different."

"How?" I asked.

He squeezed my shoulder. "By being you."

I resolved to learn to be me quickly, to please my maker and to make things different. That sounded exciting. I was, as I said, very young then.

The hardest thing, he told me, was making me fool animals. Dogs were the key. A politician without a dog was running a risk. It was an easy way to humanize them, the polls said.

That sort of thing makes me laugh.

He spent months training me — when to laugh, when to chuckle politely, when to smile with my whole face, when to merely peel my lips back. What a car-burettor is. The difference between a change purse and a dime bag. Real humans knew trillions of tiny details. I cultivated a reserved, sheltered personality for camouflage.

My maker gave me doe-like eyes and a willowy frame to further this impression. He always thought of things like that.

He managed to get an invitation to a congressman's fundraising barbecue, a man who'd surely win his party's presidential nomination. We went together, the scientist and his adopted daughter. I gave the congressman a special smile when we shook hands, and I did not put myself forward. I spent the evening listening to his mother talk about azaleas.

We were invited back for dinner. I wore black pearls, glowing against my bronze skin: traditional and modern. When some-

one asked my ethnicity, I smiled shyly and said, "American, I guess," just as I had been taught.

The congressman's smile reached his eyes. He found excuses to talk to me all evening.

Soon we were engaged. I lived in an apartment nearby, far enough to preserve decorum, close enough that we could be together always. I went to his fundraisers, his town hall meetings, his intimate chats with moneyed friends. I learned to play cribbage with his father and to understand



the racing cars his brother loved. I smiled shyly for the cameras when they caught us walking hand-in-hand down tree-lined lanes.

I learned about trade deficits, disease-resistant crops, copyright protests and prison reform.

In the stories, women like me say that they do not know how to love, and then someone comes along and teaches them. This did not happen. I did not fall in love with my congressman.

Worse. I liked him.

This made my maker's plan to destroy him problematic.

My kind were supposed to be jerky dolls, bits of plastic and metal. We were not supposed to be seamless blends. We were not supposed to fit into their families. We were not supposed to fit into their beds, not without shame and mockery. And there I was.

I did not like being a time bomb. By the

time I found out, I wasn't sure what I could do to stop it.

One day a cyberneticist came to the congressman's office. He had won an award. The staff took photos — hometown boy, making nice. They liked each other, though I couldn't tell if the cyberneticist had voted for him. Usually I could tell that sort of thing.

There was a bigger problem.

The cyberneticist knew what I was. He began to invite himself to the congressman's fundraisers. He watched me in the garden, serving lemonade to the party functionaries. He kept an eye on me at the school event, reading to children.

He knew.

One day he cornered me. "Did you think I'd let this go?"

"Let what go?" I said, feigning a pretty confusion.

"If you think I'll let you follow him into the White House to do — whatever it is you want to do — you're wrong."

"I don't know what you mean," I said. He shook his head in disgust.

I sent a message through the path I'd been given, not untraceable but hard to trace. It said: "Someone knows. He will tell the press."

The reply came back: "Good."

"I don't want this," I wrote.

The reply: "Follow the plan. All will be well."

I destroyed it automatically — that was part of the plan — but angrily — which was not. How could he say all would be well? He had not got to know the congressman. He knew his own politics.

He did not know mine.

I disappeared without a trace. No one ever knew who I was. No one ever knew who made me.

The congressman announced his candidacy that morning. He also announced a reward for any information leading to his fiancée. He intended to cooperate fully with the police in their investigation.

Those words never bode well for an election.

I went home to my maker and locked him in the basement. I feed him. Someday I will even let him out. When he's ready to hear me. When I am not his property any more. When he can think clearly again.

If that ever happens. ■

Marissa Lingen is a freelance writer living in Minnesota.

JACEY